

## Pemutakhiran Zona Musim (ZOM) Provinsi Aceh Menggunakan Data *Blending* Berbasis Non-Hirarki *K-Means Clustering*

Muhajir<sup>1,2,\*</sup>, Nazli Ismail<sup>1</sup>, Saumi Syahreza<sup>1</sup>, Andrian V H Simanjuntak<sup>2,\*</sup>

<sup>1</sup>) Departemen Fisika, Universitas Syiah Kuala (Unsyiah), Banda Aceh, Aceh

<sup>2</sup>) Badan Meteorologi Klimatologi dan Geofisika (BMKG), Aceh Besar, Aceh

Email korespondensi: muhajir1909@gmail.com; andreansimanjuntak@gmail.com

DOI: <https://doi.org/10.20527/flux.v18i1.8746>

Submitted: 09 Juli 2020; Accepted: 19 September 2020

**ABSTRAK-** Secara klimatologis, Aceh dipengaruhi oleh fenomena global, regional, dan lokal yang sangat dipengaruhi oleh topografi pegunungan berbukit. Aceh juga dikelilingi oleh Samudera Hindia di sebelah barat, Selat Malaka di sebelah timur dan Laut Andaman di sebelah utara yang membuat kondisi iklim semakin bervariasi. Zona musim lama tidak cocok untuk menjelaskan kondisi ini. Makalah ini membuat analisis terkini dengan menggabungkan data observasi dan satelit untuk penjelasan aktivitas curah hujan. Kami menggunakan k-means clustering berdasarkan analisis pencampuran data antara data observasi dan pencitraan satelit yang menghasilkan zona musim (ZOM) yang lebih spesifik dan terupdate. Cluster yang diperoleh merupakan zona musim baru yang memberikan pemetaan yang lebih spesifik untuk informasi musim kemarau dan hujan. Zona musim yang diperbarui dari 15 kluster ZOM berdasarkan k-means dapat menggambarkan variasi yang jelas sehingga akurasi perbedaan nilai curah hujan dapat lebih tepat. Dari hasil penelitian ini, zona musim yang diperbarui dapat memberikan perencanaan yang efektif tentang penguatan dari berbagai sektor pemerintahan di masa mendatang yang dapat memberikan kebijakan yang lebih baik kepada masyarakat umum.

**KATA KUNCI:** *Blending*; gugus; k-means; presipitasi; zona musim (ZOM)

**ABSTRACT-** Climatologically, Aceh is influenced by global, regional, and local phenomena strongly influenced by the hilly mountain topography. Aceh is also surrounded by Indian Ocean in the west, Malacca Strait in the east and Andaman Sea in the north that making climate conditions more varied. The old season zone is not compatible to explain this condition. This paper makes an updated analysis by combining the observational and satellite data for rainfall activity explanation. We use k-means clustering based on the analysis of blending data between observation data and satellite imaging that produces more specific and updated season zone (ZOM). The clusters obtained are newly season zones that provide more specific mapping for dry and rainy season information. The updated season zone of 15 ZOM clusters based on k-means can figure a clear variation that means the accuracy of difference rainfall value can be more precise. From this research, the updated season zone can provide effective planning about strengthened from various government sectors in the future that can give a better policy to the public community.

**KEYWORDS :** *Blending*; cluster; k-means; precipitation; season zone (ZOM)

### PENDAHULUAN

Secara klimatologis, iklim di Aceh dipengaruhi fenomena global, regional, dan lokal seperti kondisi topografi yang berbukit (Simanjuntak, Muksin & Rahmayani, 2018; Sellers & Gray, 2019). Wilayah Aceh berada di antara Samudera Hindia pada bagian barat, Selat Malaka di bagian timur dan Laut

Andaman di utara, sehingga pola curah hujan yang terjadi di wilayah Aceh menjadi sangat bervariasi (Rahman, Yang, & Di, 2018; Saha & Mitra, 2019; Edwards et al., 2020).

Secara umum, Aceh memiliki dua tipe hujan yaitu tipe *equatorial* dan tipe monsunial (Aldrian, 2003). Aceh juga terbagi ke dalam dua wilayah iklim, yaitu wilayah Zona Musim

(ZOM) dan Non-Zona Musim (Non-ZOM) seperti pada Gambar 1. Kondisi tersebut mengakibatkan variasi curah hujan dan memerlukan studi pemutakhiran dengan melakukan *clustering* (Yamanaka, 2016)



Gambar 1 Peta topografi Aceh dengan zona musim dan non-zon (N) yang lama sebanyak 5 zona, sehingga perlu dimutakhirkan.

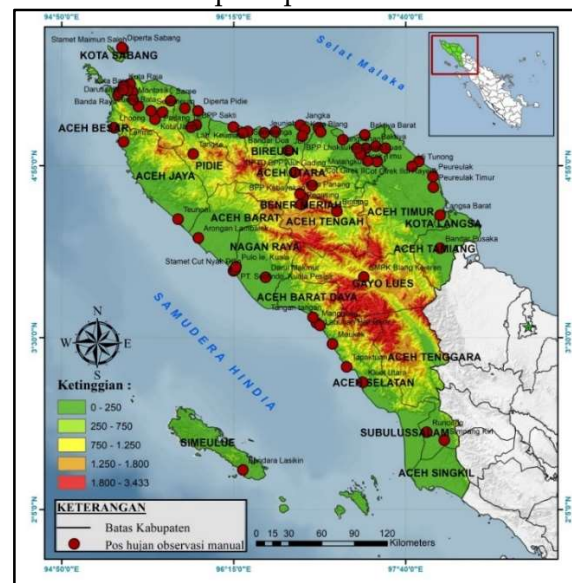
Dalam penelitian ini, penulis melakukan pemutakhiran ZOM menggunakan *blending* (pencampuran) data antara curah hujan observasi manual dengan curah hujan satelit GSMaP (*Global Satellite Mapping of Precipitation*), kemudian diklasifikasikan dengan metode Non-Hirarki. Hasil *clustering* ZOM digunakan untuk penentuan awal musim hujan (AMH) dan awal musim kemarau (AMK) di Aceh (Moron, Robertson & Pai, 2017; McGregor, 2019). Faktor lain yaitu pemanasan global yang mengakibatkan perubahan iklim sangat cepat disertai cuaca ekstrem semakin sering terjadi. Sehingga, informasi AMH dan AMK yang akurat sangat dibutuhkan. Oleh karena itu, evaluasi dan pemutakhiran ZOM di Aceh perlu dilakukan menggunakan data curah hujan rata-rata 2010-2019.

Selain itu, penelitian ini juga melakukan kajian pemutakhiran Zona Musim (ZOM) dengan menggunakan data *blending* (curah hujan observasi dan satelit) berbasis *clustering*. Dari penelitian ini, ZOM dapat lebih diperbarui, sehingga hasil prakiraan musim hujan maupun kemarau setiap wilayah ZOM

akan lebih tepat dan akurat. Dengan demikian, perencanaan dari berbagai sektor ke depan dapat dilakukan dengan lebih baik.

## METODE PENELITIAN

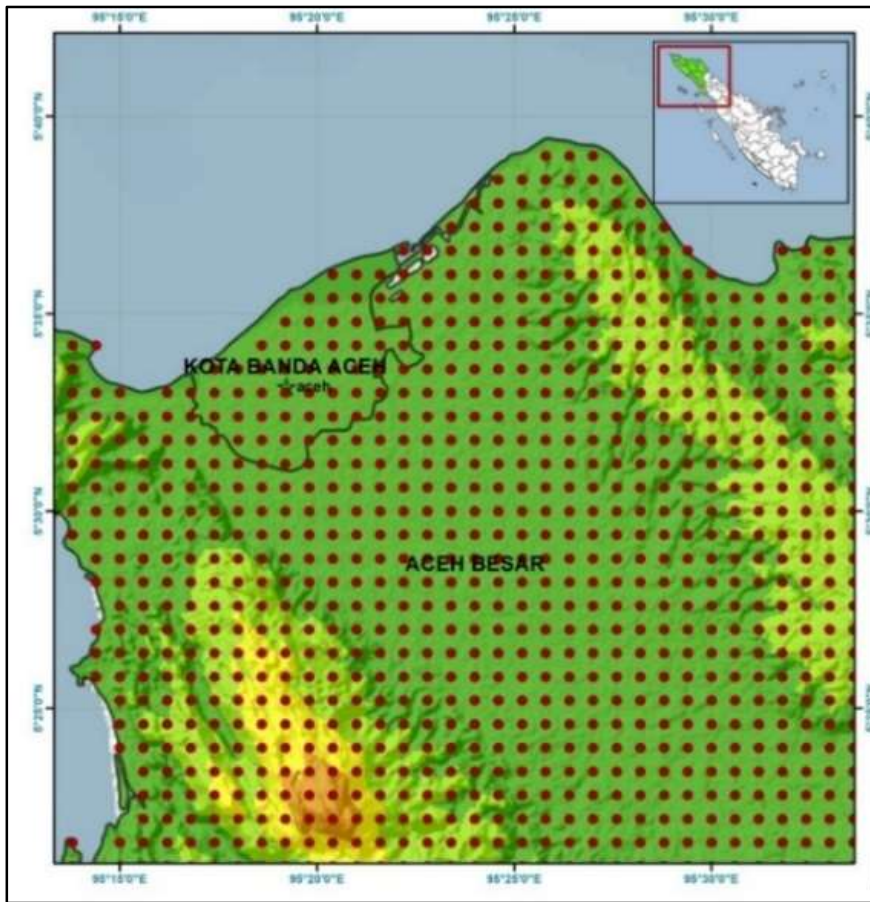
Secara geografis, Aceh terletak pada  $2^{\circ}$  -  $6^{\circ}$  LU dan  $94^{\circ}$  -  $98^{\circ}$  BT dengan iklim yang bervariasi. Untuk menjelaskan hal tersebut, iklim bisa di-*cluster* secara non-hirarki dengan data observasi seperti pada Gambar 2.



Gambar 2 Peta lokasi penggunaan data pos-pos hujan observasi di Aceh. Kompilasi data dilakukan pada sebanyak 78 pos hujan di Aceh. Data observasi digunakan untuk menganalisis curah hujan (CH) secara spatio-temporal.

Data curah hujan secara *real-field* diperlukan untuk melakukan prakiraan cuaca setiap dasarian. Selain itu, analisis curah hujan di Aceh digunakan untuk mitigasi banjir. Sehingga, data curah hujan sangat perlu dikaji (McGregor, 2019; Selers & Gray, 2019). Pada penelitian ini, kajian curah hujan dilakukan lebih komprehensif, yaitu dengan melakukan *blending* antara data lapangan dan data satelit seperti pada Gambar 3.

*Blending* dilakukan dengan menggabungkan data observasi dan satelit. Data satelit diperoleh dari GSMaP (*Global Satellite Mapping of Precipitation*) yaitu satelit meteorologi untuk pengukuran presipitasi (Ushio et al., 2004). *Blending* perlu untuk meminimalisir daerah yang secara spasial memiliki deviasi besar karena ketidakadaan pos-pos hujan (Chanda & Maity, 2016).



Gambar 3 Data pengukuran yang memiliki waktu yang sama dengan observasi diperoleh dari satelit GSMaP (*Global Satellite Mapping of Precipitation*) dengan resolusi grid 1 km x 1 km.

**Analisis Cluster**

Analisis *cluster* pada penelitian dilakukan untuk mengelompokkan obyek (curah hujan) ke dalam beberapa kelompok (ZOM) yang mempunyai sifat berbeda antara kelompok satu dengan yang lain (Pitri, Soleh, & Djuraidah, 2018). Selain itu, pengelompokan obyek dilakukan atas dasar kesamaan karakteristik. Pada penelitian ini, analisis *cluster* menggunakan metode non-hirarki memakai perhitungan *k-means* (Camus et al., 2011).

Asumsi awal analisis *cluster* sebelum dilakukan *k-means* yaitu meminimalisir *outlier* agar sampel yang digunakan bisa mewakili. Pengujian sampel dilakukan dengan uji *Bartlett* dengan hipotesis  $H_0$  (sampel belum layak) dan  $H_1$  (sampel sudah layak) dan uji *Kaiser Mayer Olkin* (KMO).

$$B = \frac{(\sum_{i=1}^p v_i) \ln (\sum_{i=1}^p v_i S_i^2) - (\sum_{i=1}^p v_i \ln S_i^2)}{1 + \{\sum_{i=1}^p (v_i \ln S_i^2)\}} \quad (1)$$

dengan :  $S_i^2 = \frac{\sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2}{(n_i - 1)}$  (2)

$P$  = n-variabel

$V$  = derajat bebas

$V_i = n_i - 1$

$N$  = jumlah observasi/ obyek ke  $- i$

$B$  = uji statistik *Bartlett*

Kriteria Pengambilan Keputusan yaitu,  $H_0$  ditolak jika  $B > X_{a,p(p-1)/2}^2$  atau probabilitas dibagi nilai  $sig < a$ , dan bisa dilakukan uji *Kaiser-Mayer-Olkin* (KMO)

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p p_{ij}^2} \quad (3)$$

dengan :

$r_i$  = korelasi variabel  $i$  dan  $j$  ( $i < j$ )

$p_{ij}$  = korelasi parsial variabel  $i$  dan  $j$  ( $i < j$ )

Jika,  $r_i = 0$ , variabel pengukuran merupakan sebuah faktor dan  $KMO = 1$ .

Jika,  $p_{ij} = 1$ , variabel pengukuran bukan sebuah faktor dan  $KMO = 0$ .

Jika nilai  $KMO \geq 0.5$ , maka sampel sudah layak untuk dianalisis.

**Non-Hirarki Cluster**

Metode non-hirarki bekerja dengan menentukan terlebih dahulu jumlah *cluster*



yang diinginkan (dua, tiga, atau yang lain). *Clustering* non-hierarki atau *K-Means clustering* menggunakan iterasi untuk memilih pusat (*centroid*) yang bersifat sementara dengan terus diperbaharui sampai kriteria tercapai (Kappe, Bottinger, & Lette, 2019; Straus, 2019).

Metode *k-means* bersifat tanpa arahan, karena analisis dalam pengelompokan tidak mempunyai *cluster* yang pasti. Objek yang sudah masuk ke dalam *cluster* tertentu masih bisa berpindah ke *cluster* yang lain. Strauss (2019) berpendapat bahwa istilah *k-means* lebih cocok untuk menandai setiap objek masuk ke dalam *cluster* yang mempunyai pusat *cluster* (rata-rata) terdekat.

$$K - \text{Means } (J) = \sum_{i=1}^m \sum_{k=1}^K w_{ik} \|x^i - \mu_k\|^2 \quad (4)$$

dimana,

$$w = \begin{cases} 1 & \text{jika } k = \text{argmin}_j \|x^i - \mu_k\|^2 \\ 0 & \text{jika } k \text{ tidak memiliki klaster} \end{cases}$$

Terdapat dua analisis mendasar dalam mencari solusi non-hirarki. Pertama, solusi harus berangkat dari banyak *cluster* yang sebelumnya ditentukan terlebih dahulu. Kedua, pemilihan *centroid cluster* tidak begitu bisa diyakini, sehingga peng-*clusteran* bergantung pada total *centroid* yang terpilih (Kappe, Bottinger, & Lette, 2019; Zscheischler, Mahecha, & Harmeling, 2012).

Banyak program yang dimulai dengan memilih kasus pertama  $k$  ( $k$  = jumlah *cluster*) sebagai pusat *cluster* awal. Jadi, hasil *cluster* tergantung pada observasi data. Dibalik kekurangan itu, metode ini dapat dilakukan dengan cepat dan sangat bermanfaat untuk jumlah data yang besar (Netzel & Stepinski, 2016; Thiede & Gray 2017; Pitri, Soleh, & Djuraidah, 2018).

## HASIL DAN PEMBAHASAN

Hasil penelitian ini menunjukkan suatu kondisi yang jelas pada perbedaan data manual dan satelit dari hasil interpolasi. Selain itu, penggabungan keduanya bertujuan memperbaiki model ZOM sebelumnya dan memberikan gambaran yang komprehensif.

### Curah Hujan *Real-Field*

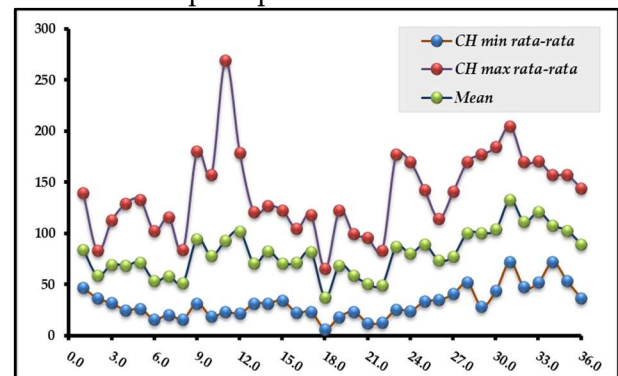
Curah hujan lapangan dari observasi

dianalisis secara spasial. Curah hujan *real-field* seperti pada Gambar 4 merupakan hasil observasi untuk kepentingan prakiraan.



Gambar 4. Peta analisis spasial curah hujan dari observasi lapangan pada dasarian III di bulan Oktober 2019. Analisis observasi lapangan sangat penting untuk melihat tren suatu wilayah dan selanjutnya di-*blending* dengan data satelit.

Curah hujan observasi dikompilasi dari pos-pos hujan di seluruh Aceh seperti pada Gambar 4. Untuk menghasilkan zona musim yang akurat dan sesuai, data digabung dengan data satelit seperti pada Gambar 5.

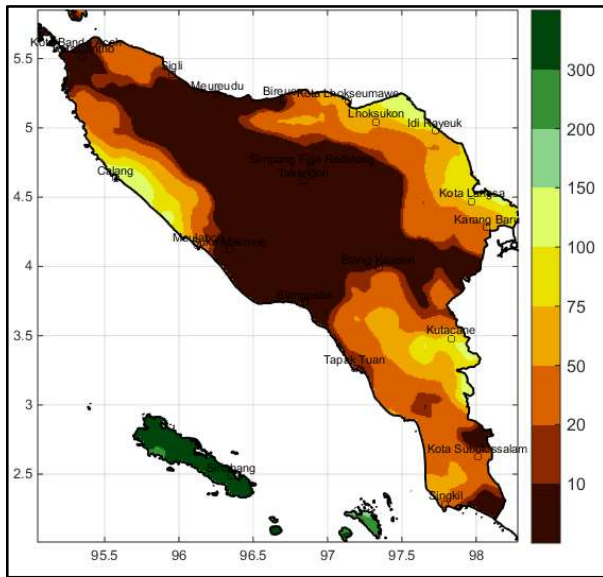


Gambar 5 Grafik temporal curah hujan (CH) untuk total data 10 tahun (2010 – 2019). Dasarian pada grafik mengacu pada total analisis 10 hari, sehingga, satu tahun memiliki sebanyak 36 dasarian dengan nilai-nilai yang berbeda.

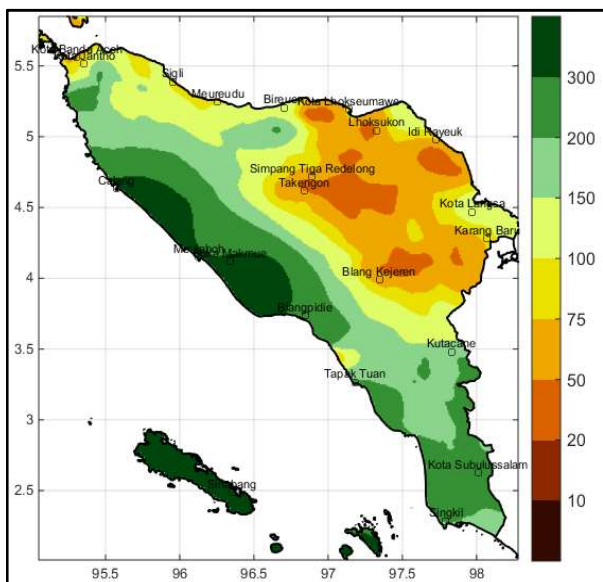
### Curah Hujan Satelit *GSMaP*

Analisis data curah hujan (CH) satelit *GSMaP* dianalisis secara spasial dengan interpolasi  $0.01^\circ$  (1 km) seperti pada Gambar 6.

Data pengukuran GSMaP digunakan pada penelitian ini untuk analisis data jangka panjang zona musim (ZOM).



Gambar 6 Peta data curah hujan (CH) dalam mm, dari pengukuran satelit GSMaP pada dasarian yang sama di Oktober 2019. Data satelit memiliki resolusi dengan grid yang cukup rapat sehingga sangat baik untuk di-*blending* dengan daerah yang tidak memiliki data observasi.



Gambar 7 Data *blending* observasi dan satelit dengan grid 1 km. Data komprehensif dan *dense* bertujuan untuk melengkapi ketiadaan observasi dari daerah yang tidak memiliki pos-pos pengamatan hujan. Gambar di atas merupakan hasil interpolasi pada dasarian III Oktober 2019.

## Zona Musim (ZOM)

Selanjutnya, data observasi dan satelit di-*blending*. Hasil *blending* merupakan gabungan dari observasi dan satelit. Hasil *blending* memberikan informasi spasial yang lebih baik dan komprehensif, karena bisa meng-*cover* daerah-daerah yang tidak ada pos hujan seperti pada Gambar 7. Hasil analisis spasial secara interpolasi pada data *blending* menunjukkan beberapa wilayah Aceh memiliki variasi nilai dengan deviasi kecil.

Hasil analisis pada data *blending* dengan *k-means* menghasilkan sebanyak 15 *cluster* zona musim (ZOM). ZOM sebelumnya hanya sebanyak lima, sehingga belum cukup baik menjelaskan periode musim tiap kabupaten. Dengan pemutakhiran berbasis *k-means*, zona musim menjadi lebih akurat sehingga informasi curah hujan dan musim hujan pada tiap daerah lebih pasti dari sebelumnya.

Luas ZOM tidak pernah selalu sama dengan luas daerah administrasi, yang mana satu daerah bisa terdiri dari beberapa ZOM dan sebaliknya. Pemodelan ZOM pada penelitian ini diperbaiki secara *real-time* dengan mempertimbangkan perubahan data secara *time series* pada rentang 10 tahun. Sehingga, hasil yang didapatkan memberikan model seperti pada Gambar 8 menjadi lebih baik dan jelas dalam mengelompokkan daerah-daerah yang memiliki ZOM berbeda.

Model ZOM yang diperoleh dari data *blending* memiliki nilai deviasi yang kecil sehingga lebih dipercaya untuk menjelaskan pentingnya perbedaan zona musim yang menggambarkan variasi kondisi iklim di Aceh seperti pada Gambar 8. Zona musim seperti pada Gambar 8 diperoleh dengan meng-*cluster* sebanyak *n-grid* total penggabungan data *blending* antara manual dan satelit seperti pada Tabel 1. Data terbanyak yaitu *cluster* 6 dengan luas area musiman yang juga cukup luas dan mayoritas berada di tenggara Aceh. Sedangkan, data paling sedikit yaitu *cluster* 11 yang letaknya di Aceh bagian tengah.





## KESIMPULAN

Analisis berbasis *k-means clustering* pada data *blending* antara data curah hujan observasi dan satelit mampu menghasilkan *cluster* yang lebih spesifik. *Cluster* merupakan zona musim yang memberikan informasi musim kemarau dan hujan lebih spesifik. Zona musim yang dimutakhirkan sebanyak 15 *cluster* berbasis *k-means* dari yang sebelumnya lima *cluster*. Pemutakhiran ini dapat menginterpretasikan perbedaan yang jelas, sehingga keakuratan perbedaan curah hujan (CH) di setiap tempat bisa lebih presisi.

## DAFTAR PUSTAKA

- Aldrian, E. (2008). Meteorologi Laut Indonesia. Jakarta: Badan Meteorologi dan Geofisika.
- Ayundari, I., & Sutikno, S. (2019). Penentuan Zona Musim di Mojokerto Menurut Karakteristik Curah Hujan Dengan Metode Time Series Based Clustering. *Inferensi*, 2(2), 63-70.
- Camus, P., Mendez, F. J., Medina, R., & Cofiño, A. S. (2011). Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coastal Engineering*, 58(6), 453-462.
- Chanda, K., & Maity, R. (2016). Uncovering global climate fields causing local precipitation extremes. *Hydrological Sciences Journal*, 61(7), 1227-1237.
- Edwards, R. B., Naylor, R. L., Higgins, M. M., & Falcon, W. P. (2020). Causes of Indonesia's forest fires. *World Development*, 127, 104717.
- Kappe, C. P., Böttinger, M., & Leitte, H. (2019, June). Analysis of Decadal Climate Predictions with User-guided Hierarchical Ensemble Clustering. *In Computer Graphics Forum*, 38(3), 505-515.
- McGregor, G. R. (2019). Climate and rivers. *River Research and Applications*, 35(8), 1119-1140.
- Moron, V., Robertson, A. W., & Pai, D. S. (2017). On the spatial coherence of sub-seasonal to seasonal Indian rainfall anomalies. *Climate Dynamics*, 49(9-10), 3403-3423.
- Netzel, P., & Stepinski, T. (2016). On using a clustering approach for global climate classification. *Journal of Climate*, 29(9), 3387-3401.
- Pitri, R., Soleh, A. M., & Djuraidah, A. (2018). Statistical downscaling modeling through k-means clustering. *International Journal of Scientific Research in Science, Engineering and Technology*, 4(9), 220-227.
- Rahman, M., Yang, R., & Di, L. (2018). Clustering Indian Ocean tropical cyclone tracks by the standard deviational ellipse. *Climate*, 6(2), 39.
- Saha, M., & Mitra, P. (2019). Identification of Indian monsoon predictors using climate network and density-based spatial clustering. *Meteorology and Atmospheric Physics*, 131(5), 1301-1314.
- Sellers, S., & Gray, C. (2019). Climate shocks constrain human fertility in Indonesia. *World Development*, 117, 357-369.
- Simanjuntak, A. V., Muksin, U., & Rahmayani, F. (2018). Microtremor survey to investigate seismic vulnerability around the Seulimum Fault, Aceh Besar-Indonesia. *MS&E*, 352(1), 012046.
- Straus, D. M. (2019). Clustering Techniques in Climate Analysis. In *Oxford Research Encyclopedia of Climate Science*.
- Thiede, B. C., & Gray, C. L. (2017). Heterogeneous climate effects on human migration in Indonesia. *Population and Environment*, 39(2), 147-172.
- Ushio, T., Okamoto, K. I., Iguchi, T., Takahashi, N., Iwanami, K., Aonashi, K., Shige, S., Hashizume, H., Kubota, T., & Inoue, T. (2003). The global satellite mapping of precipitation (GSMaP) project. *Aqua (AMSR-E)*, 2004.
- Yamanaka, M. D. (2016). Physical climatology of Indonesian maritime continent: An outline to comprehend observational studies. *Atmospheric Research*, 178, 231-259.
- Zscheischler, J., Mahecha, M. D., & Harmeling, S. (2012). Climate classifications: the value of unsupervised clustering. *Procedia Computer Science*, 9, 897-906.