**Extended-range prediction of monsoon onset over Northern Vietnam and the Gulf of Tokin**

Tuan Bui-Minh1 and Kim-Cuong Nguyen1\*

1VNU University of Science, Vietnam National University, Hanoi, Vietnam

Corresponding author address: University of Science, Vietnam National University, 334 Nguyen Trai Street, Thanh Xuan, Hanoi, 100000 Vietnam

Email: Kim-Cuong Nguyen, [cuongnk@hus.edu.vn](mailto:cuongnk@hus.edu.vn)

**Abstract**

The onset of the rainy season in Vietnam displays distinct characteristics compared to other Asian summer monsoon (ASM) regions due to its elongated shape running from south to north, coupled with its complex terrain and influences of multiple weather systems. Determining the onset is challenging, although numerous empirical indices exist. Here, we proposed an alternative approach to determine the onset by examining the large-scale weather patterns associated with the abrupt increase in rainfall in the country. In this new approach, the evolutions of the ASM system were interpreted as consecutive occurrences of multiple weather patterns, detected by a clustering algorithm Self Organizing Maps (SOMs). Specifically, the SOMs were applied on outgoing longwave radiation, 850 hPa wind, and geopotential height over the ASM regions from April to July. The primary features of these atmospheric variables were generalized to 16 patterns. A high-resolution Vietnam Gridded Precipitation (VnGP) dataset was employed to investigate the linkage between these large-scale patterns and local-scale rainfall. The large-scale patterns associated with the onset were distinguished, which are characterized by the development of westerly winds and large-scale convections. Then, the onset of summer rainfall in Vietnam is identified based on the initial appearance of wet-condition weather patterns. This approach offers several advantages: it is objective and does not require additional thresholds of rain or wind to determine the monsoon onset. It also more effectively captures the characteristics of large-scale circulation and deep convection compared to relying on predefined monsoon indices alone. Additionally, this approach helped to elucidate two types of weather patterns associated with the onset: those associated with the northeastward development of deep convection and westerly wind from the Indian Ocean to the Indochina Peninsula and activities of the Boreal Summer Intraseasonal Oscillation.

**Keywords**: rainy season onset, intraseasonal oscillation, tropical disturbance, self-organizing maps, monsoon index, Foehn effect

1. **Introduction**

Northern Vietnam and the Gulf of Tonkin (Fig. 1) are two key drivers of Vietnam economic growth, leveraging their strategic geographic location, industrial development, and expanding infrastructure. Northern Vietnam is a hub for manufacturing, particularly in electronics, textiles, and automotive industries. It is also a leading producer of agricultural goods, particularly rice and tea, which contribute to both domestic supply and export markets. Located to the east of Northern Vietnam, the Gulf of Tonkin plays a crucial role in sectors such as trade, fishing, and energy. Bordering both Vietnam and China, the gulf serves as a significant maritime route, facilitating trade between these two countries and the broader Southeast Asian region. The Gulf is also rich in marine resources, supporting a thriving fishing industry that contributes to the livelihoods of local communities.

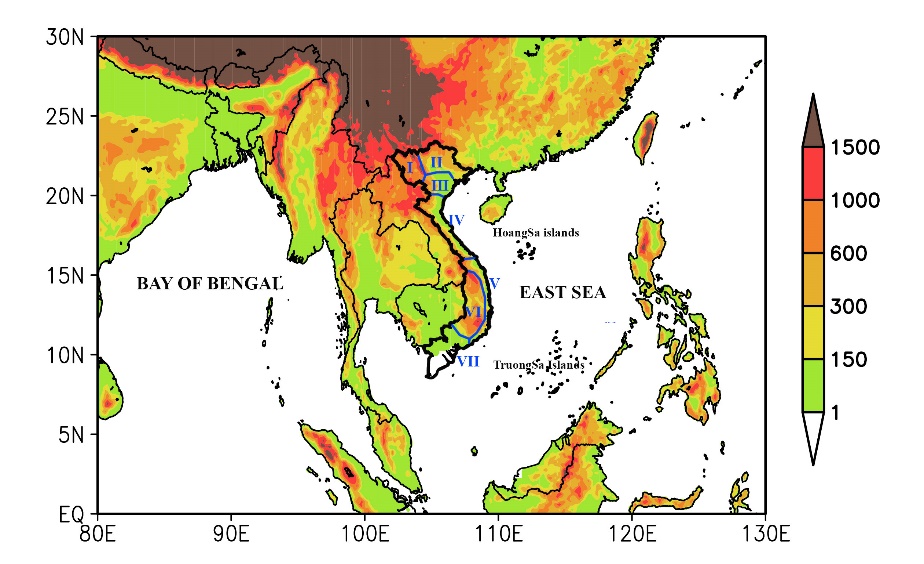


Figure 1. Northern Vietnam and the Gulf of Tonkin

The onset of summer monsoon leading to the increase of rainfall, which is important to the socio-economic well-being of North Vietnam and the Gulf of Tokin. Additionally, changes in surface winds drive alterations in ocean currents, leading to significant environmental and ecological impacts, including shifts in nutrient transport, marine productivity, fisheries, and pollution dispersal. Therefore, identifying and predicting the monsoon onset is essential for the subregions. However, duo to the complex chracteristics of rainfall and circulation, these tasks is still a topic of controversial (Kiguchi et al., 2016; Nguyen-Le et al. 2014; Bombardi et al. 2021; Bui-Minh et al. 2024).

Traditional studies used two type of monsoon indices to identify monsoon onset. The first are dynamic indices, which are constructed by transforming large-scale signals, such as the reversal of low-level wind or upper-level temperature gradient into time series. The second are based on the abrupt increase of rainfall, with the concept of monsoon as an alternation between dry and wet season. Although these two indices have physical meaning, they display many disadvantage in identifying the monsoon onset in Northern Vietnam. For more detailed, Zhang et al., (2002) pointed out that there are at least three weather systems that simultaneously affect the subregions, that is, the southwesterly monsoon wind, mid-latitude westerly winds from northern India and southwesterly winds on the flank of the Western North Pacific Subtropical High (WNPSH). Since all these systems involve westerly winds, using tropical westerly wind or low-level wind reversal as an indicator of monsoon onset is problematic (Wang et al. 2004; Kiguchi et al. 2016; Ngo-Thanh et al. 2018; Pham-Xuan et al. 2010). Meanwhile, rainfall often occurs earlier than the occurrence of the tropical westerlies, especially in the Northern Vietnam due to influence of cold surges and tropical disturbances (Nguyen et al. 2022; Nguyen-Le et al. 2014), making difficult to define monsoon onset as the start of rainy season.

Most study attempt to predict the monsoon onset over the Southern Vietnam. (Pham et al. 2010; Pham-Thanh et al. 2019).

Recently, Bui-Minh et al., (2024a, b) proposed a novel method for identifying the monsoon onset in Vietnam through weather-pattern recognition. This approach interprets the evolution of atmospheric systems as a sequence of patterns, with the monsoon onset marked by the first appearance of specific weather patterns. These onset patterns are characterized by the onset of deep convection and the development of monsoon westerlies. Additionally, these patterns must frequently occur during summer to ensure the persistence of monsoon rainfall and circulation. Because the method combines large-scale circulation and deep convection patterns, it can distinguish between pre-monsoon and monsoon rainfall, a challenge for traditional methods. Furthermore, it provides a better explanation for the variation in onset dates across different locations compared to methods based solely on subjective thresholds of monsoon indices.

This synoptic approach offers an alternative way to predict the monsoon onset over North Vietnam and the Gulf of Tonkin. Instead of predicting the monsoon onset based on local rainfall or predefined monsoon indices which are highly sensitive to the changes of domain choice, variables and criteria, numerical models can be used as a guide to predict a finite number of weather-pattern clusters related to the onset. The first advantage is the large-scale patterns are often predictable, even at medium and extended ranges.

(Bombardi et al. 2017; 2021; Alessandri et al. 2015; Cherchi and Navarra, 2003; Li and Zhang, 2009).

Section 2 describes the data and methodologies used in this study. The evolutions of large-scale pattern and their relationships with the onset of monsoon rainfall in Vietnam are documented in Section 3. Finally, discussion and conclusions are given in Sections 4 and 5.

1. **Data and methodology**

*2.1. Data*

This study's primary dataset comes from the NCEP-DOE reanalysis 2, providing daily mean meteorological fields. This dataset was sourced from NOAA/OAR/ESRL PSL in Boulder, Colorado, USA (Kanamitsu et al. 2002). Other significant data is the daily Outgoing Longwave Radiation (OLR) data obtained from NOAA satellites (Liebmann and Smith 1996). The OLR served as a proxy for investigating the activities of convections in tropical and subtropical regions. The reanalysis and OLR data are at 2.5° × 2.5° latitude-longitude resolution, providing information on the atmospheric processes at a large scale. The detailed variations of Vietnam rainfall are analyzed based on the high-resolution rainfall data VnGP (Nguyen-Xuan et al. 2016) at a resolution of 0.1° × 0.1° latitude-longitude. The analysis encompasses data from 1981 to 2010.

*2.2. Methodology*

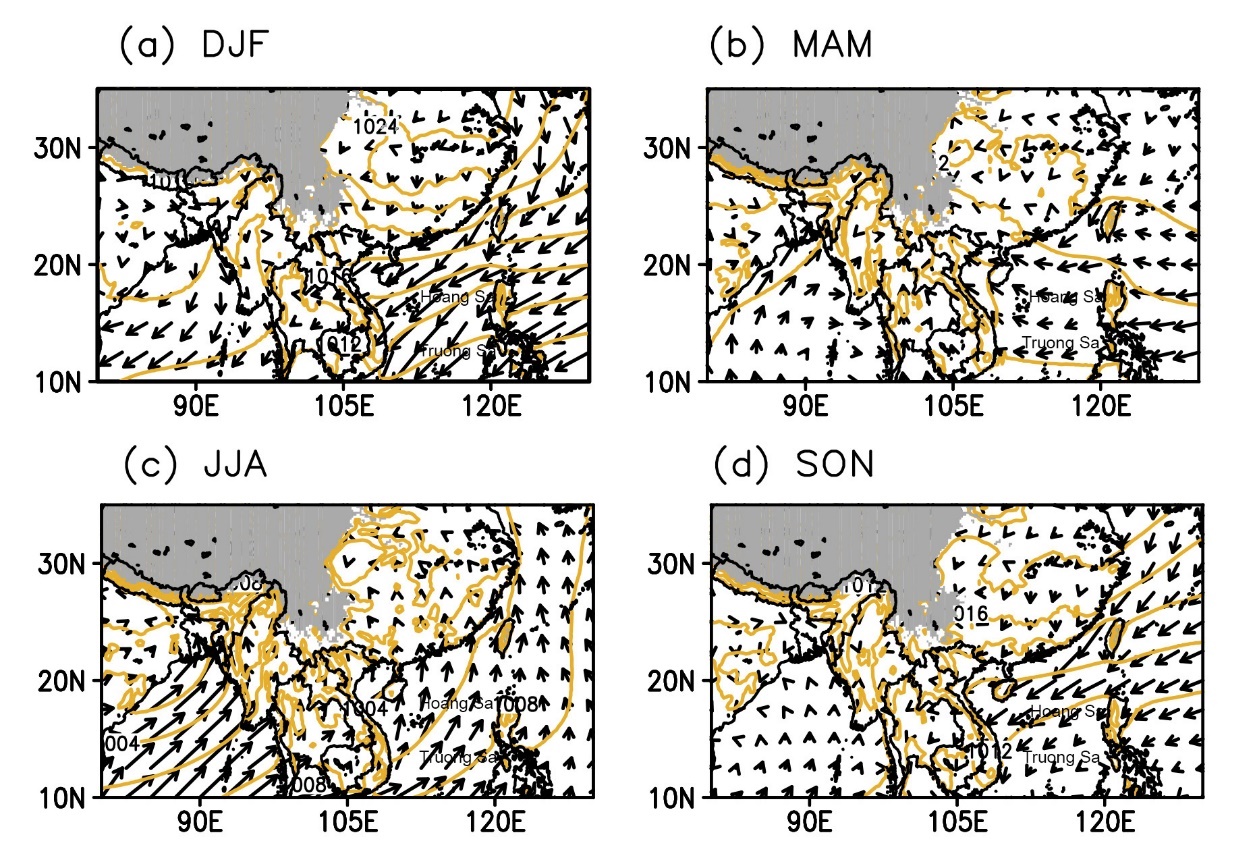
Self-Organizing Maps (SOMs) were applied to the dataset to classify early summer weather patterns into a 10x10 nodes, which were then organized into 6 clusters by k-means. The monsoon onset were identify as the initial occurrence of weather patterns belonging the the monsoon clusters.

Once the nodes were defined, the squared Euclidean distance between the raw S2S forecast (at each lead time) and each weightnode was calculated to assess the forecast's accuracy.

1. **Result**

*3.1. Climatological mean of circulation and rainfall*

Northern Vietnam experiences four distinct seasons, each characterized by typical weather conditions. In winter, southward intrusions of the Siberian high, often referred to as cold surge, bring cold and dry weather across the entire subregion (Nguyen et al., 2014; Nguyen et al., 2022). In spring, the cold surges changes their pathways from southward to southeastward, shifting prevailing winds from northwesterly to southeasterly that carry significant moisture from the Gulf of Tonkin. In combination with low surface temperature, the winds lead to occurences of dense fog and drizzle across the Red River Delta (Nguyen et al., 2022). As summer approaches, the Siberian high weakens, the southeasterly winds are subsequently replaced by monsoonal southwesterly winds, which create favorable conditions for the development of deep convection (Zhang et al., 2004; Nguyen-Le et al., 2014). This change of low-level winds triggers an abrupt increase of rainfall over Northern Vietnam and the Gulf of Token (Bui-Minh et al. 2024; 2014; Nguyen et al., 2014). The rainfall remains substantial in the following months, reaching its peak in August. In autumn, while the westerly winds weaken and retreat, the Siberian high reasserts its influence, bringing back the cool and dry conditions to the subregions.



*3.1. Identifying monsoon onset over North Vietnam*

The mean patterns of 850 hPa wind, geopotential height and total precipitation in the six SOMs nodes in the first experiment were displayed in the Fig. 3. The order of the node is based on the distribution in frequency of the node by time (Fig. 4a). The result display a similar result with previous studies (Tuan et al. 2024a, b) that the SOMs effectively classify the patterns associated with dry and wet conditions associated with large-scale circulation over the South and Southeast Asian. For more detailed, in the node 0-0 and 1-0, anticyclonic circulation associated with the WNP high dominate over the Indochina Peninsula while convection is mainly observed over the Maritime continent or equatorial Indian Ocean. Therefore, these two nodes present the dry condition before over Northern Vietnam and the Gulf of Tokin. In a different manner, in node 0-1, westerly wind prevail over South and Southeast Asian, that is associated with the northward jump of deep convection to the Bay of Bengal, Indochina Peninsula and Northern Vietnam. In the node 1-1, 1-2, 0-2, the convection expanse further to the northeast and northwester direction. This result indicating the the node 1-1, 1-2, 0-2 present the wet condition associated with the development of westerly monsoon over Northern Vietnam and the Gulf of Tokin.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
| Figure 1. Mean patterns of 850 hPa wind, geopotential height and total precipitation in each SOMs node | |

It is important to note that, the weather patterns in the each node only occur over specific period, indicating the effectiveness of SOMs in classifying the weather patterns in each regime. This results suggest that we can identify the monsoon onset by examining the transition of weather patterns from these nodes. The node 0-0 and node 0-1 tend to occurs in earlier days (from the pentad 18 to 22) that is present the dry conditions before onset. The nodes 1-0 and 1-1 appears around pentad 22-24, that characterized by wet-conditions associated the development of westerly, that is associated with the monsoon onset. Finally, node 1-2 and node 0-2, characterized by the northward migration of deep convection, present for the atmospheric conditions after onset. It is important to note that, if monsoon is defined as the arrival of tropical westerly wind, the monsoon onset might occurs in the entire South and Southeast Asian regions. However, if monsoon is defined as the transition from dry to wet season driven by large-scale circulation, the monsoon onset earlier over some specific regions (such as the southern Bay of Bengal and Indochina Peninsular) than over other regions.

In Northern Vietnam and the Gulf of Tokin, the wind direction is nearly unchanged in the six mean patterns, we still recognize the different large-scale circulation systems influence the subregion. In the patterns of the node 0-0 and 0-1, the wind is primarily related to the activities of the WNP subtropical high. On the other hand, in the node 1-0 and 1-1, the wind synchronized with the development of tropical westerly winds and abrupt increase of rainfall, indicating the onset of monsoon over the subregions. In contrast, a given threshold of rainfall or wind over the subregions might not clearly indicating the natures of rainfall or wind, that might leads to imprecise monsoon onset date. Some additional thresholds were often include to ensure the persistent of rain or wind, such as the rainfall must be remain positive in at least three pentad after the onset. However, the choice of thresholds are subjective and not always reflect the true nature of monsoon rain and wind. This highlight the advantages of using cluster analysis in determining the monsoon onset at regional scale compared to large-scale monsoon indices.

The mean patterns of 850 hPa wind, geopotential height and total precipitation in the six SOMs nodes in the second experiments were displayed in the Fig. 4. Although the precipitation is not included in the input data for cluster analysis in the first experiment, we still can identify the days (weather patterns) belonging to the SOMs’ nodes, which can be used to calculate the mean pattern of atmospheric variables. It can be seen that, there are two type of weather patterns are recognized, that are related to the wet and dry condition over the South and Southeast Asian. However, the temporal distribution of the weather patterns in the each nodes are expands to an entire period (Fig 4b), indicating the the dry- and wet-conditions weather patterns were not effectively classified. A various of weather patterns in the latter stages, which should be classified as wet-conditions weather patterns but is classified as dry condition weather patterns due to the lack of precipitation information. It indicating that similar circulation patterns might not always associated with similar rainfall patterns. Therefore, precipitation is crucial variables which should not be omitted from the input data for cluster analysis of weather patterns.

*3.2. Prediction of the monsoon onset*

1. **Discussion**
2. **Conclusion**

**Acknowledgment**

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 105.06-2021.28.

**Conflict of interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Data Availability Statement**

The OLR is downloaded at <https://psl.noaa.gov/data/gridded/data.olrcdr.interp.html>. The NCEP\_DOE reanalysis 2 data are provided at [https://psl.noaa.gov/data/gridded/data.ncep.reanalsis2.pressure.html](https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.pressure.html). The VnGP rainfall data is available to the public via <https://search.diasjp.net/en/dataset/VnGP_010>. The BSISO index data was downloaded at <https://cliks.apcc21.org/dataset/bsiso>

**References**

1. Bombardi R. J., Moron V., Goodnight J. S., 2020: Detection, variability, and predictability of monsoon onset and withdrawal dates: A review. *Int J Climatol.* ; 40: 641–667. <https://doi.org/10.1002/joc.6264>
2. Bordoni S., Schneider T., 2018, Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nature Geosci* *1*, 515–519. <https://doi.org/10.1038/ngeo248>
3. Cassano E. N., Glisan J. M., Cassano J. J., Gutowski W. J., Seefeldt M. W., 2015: Self-organizing map analysis of widespread temperature extremes in Alaska and Canada. *Clim Res,* 62(3):199–218
4. Cassano E.N., Cassano J.J., Seefeldt M.W., Gutowski W.J., and Glisan, J.M., 2017: Synoptic conditions during summertime temperature extremes in Alaska. *Int J Climatol*, 37: 3694-3713. <https://doi.org/10.1002/joc.4949>
5. Chao W. C., and Chen B., 2001: The Origin of Monsoons. *J Environ Sci*, 58, 3497–3507
6. Chu J. E., Hameed S. N., & Ha, K. J., 2012: Nonlinear, intraseasonal phases of the East Asian summer monsoon: Extraction and analysis using self-organizing maps. *J Climate*, 25(20), 6975-6988.
7. Chu J., Wang B., Lee J., and Ha K., 2017: Boreal Summer Intraseasonal Phases Identified by Nonlinear Multivariate Empirical Orthogonal Function–Based Self-Organizing Map (ESOM) Analysis. *J Climate*, 30, 3513–3528, <https://doi.org/10.1175/JCLI-D-16-0660.1>.
8. Deme A., Viltard A., de Felice P., 2003: Daily precipitation forecasting in Dakar using the NCEP-NCAR reanalyses. *Wea Forecast* 18:93–105, <https://doi.org/10.1175/1520-0434(2003)018%3C0093:DPFIDU%3E2.0.CO;2>.
9. Ding Y., 2007: The Variability of the Asian Summer Monsoon, *J Meteorol Soc Jpn*, Ser. II, 85B, 21-54,  <https://doi.org/10.2151/jmsj.85B.21>.
10. Doan Q.-V., Kusaka, H., Sato, T., and Chen, F., 2021: S-SOM v1.0: a structural self-organizing map algorithm for weather typing, *Geosci Model Dev*, 14, 2097–2111, https://doi.org/10.5194/gmd-14-2097-2021,
11. Fasullo J., and Webster P. J., 2003: A Hydrological Definition of Indian Monsoon Onset and Withdrawal. *J Climate*, 16, 3200–3211, <https://doi.org/10.1175/1520-0442(2003)016%3C3200a:AHDOIM%3E2.0.CO;2>.
12. Gadgil S., 2018: The monsoon system: Land–sea breeze or the ITCZ? *J. Earth Syst. Sci.*, *127*(1), 1-29.
13. Geen R., Bordoni S., Battisti D. S., & Hui K., 2020: Monsoons, ITCZs, and the concept of the global monsoon. *Rev Geophys*, 58, e2020RG000700.
14. Gill A. E. (1980). Some simple solutions for heat‐induced tropical circulation. *Q J R Meteorol Soc*, 106(449), 447-462.
15. Gueye A. K., Janicot S., Niang A., Swadogo S., Sultan B., Diongue-Niang A., and Thiria S., 2011: Weather regimes over Senegal during the summer monsoon season with self-organizing maps. Part I: Synoptic time scale. *Climate Dyn*, 36, 1–18.
16. Hewitson B. C., and Crane R. G., 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res*, 22, 13–26.
17. Holton J. R., and  Colton D. E., 1972: A diagnostic study of the vorticity balance at 200 mb in the tropics during the northern summer, J Atmos Sci, **29,**1124–1128.
18. Islam M. R., Sheridan S. C., & Lee C. C., 2019: Using self-organizing maps to identify the South Asian seasonal cycle, *Theor Appl Climatol*, 137, 1385-1401.
19. Johnson N. C., 2013: How Many ENSO Flavors Can We Distinguish?. *J Climate*, 26, 4816–4827, <https://doi.org/10.1175/JCLI-D-12-00649.1>.
20. Kanamitsu M., Ebisuzaki W., Woollen J., Yang S. K., Hnilo J. J., Fiorino M., and Potter G. L., 2002: NCEP–DOE AMIP-II reanalysis (R-2). *Bull Am Meterol Soc*, *83*(11), 1631-1644.
21. Kikuchi K., 2021: The boreal summer intraseasonal oscillation (BSISO): A review. *JMSJ Ser II*, 99(4), 933-972, <https://doi.org/10.2151/jmsj.2021-045>.
22. Kohonen T. (1990): The Self-Organizing Map. Proceedings of the IEEE, 78, 1464-1480. <http://dx.doi.org/10.1109/5.58325>
23. Liebmann B., & Smith C. A. 1996: Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull Am Meterol Soc*, 77(6), 1275-1277.
24. Luong T. M., Dasari H. P., Doan Q., Alduwais A. K., and Hoteit I. (2024): Organized precipitation and associated large-scale circulation patterns over the Kingdom of Saudi Arabia. *Int J Climatol* Portico. <https://doi.org/10.1002/joc.8524>.
25. Matsumoto J., 1997: Seasonal Transition of Summer Rainy Season over Indochina and Adjacent Monsoon Region. *J Adv Atmos Sci*, 14(2): 231, <https://doi.org/10.1007/s00376-997-0022-0>.
26. Neelin J., 2007: Moist dynamics of tropical convection zones in monsoons, teleconnections, and global warming. *The Global Circulation of the Atmosphere*, T. Schneider and A. H. Sobel, Eds., Princeton University Press, 267–301.
27. Ngo-Thanh H., Ngo-Duc T., Nguyen-Hong H. et al., 2018: A distinction between summer rainy season and summer monsoon season over the Central Highlands of Vietnam. *Theor Appl Climatol*, **132**, 1237–1246, <https://doi.org/10.1007/s00704-017-2178-6>.
28. Nguyen-Le D., Matsumoto J., and Ngo-Duc T., 2015: Onset of the Rainy Seasons in the Eastern Indochina Peninsula.*J Climate*, **28**, 5645–5666, <https://doi.org/10.1175/JCLI-D-14-00373.1>
29. Nguyen-Le D., Matsumoto J., and Ngo-Duc T., 2014: Climatological onset date of summer monsoon in Vietnam. Int J Climatol, **34**, 3237–3250, <https://doi.org/10.1002/joc.3908>
30. Nguyen-Xuan T., Ngo-Duc T., Kamimera H., Trinh-Tuan L., Matsumoto J., Inoue T., and Phan-Van T., 2016: The Vietnam Gridded Precipitation (VnGP) Dataset: Construction and validation. *SOLA*, 12, 291–296, <https://doi.org/10.2151/sola.2016-057>.
31. Ohba M., Nohara D., and S. Kadokura, 2016: Impacts of Synoptic Circulation Patterns on Wind Power Ramp Events in East Japan, *J Renew Energy*, 96, 591–602. doi:10.1016/j.renene.2016.05.032
32. Ohba M., Nohara D., Kadokura S., and Toyoda Y., 2016: Rainfall Downscaling of Weekly Ensemble Forecasts using Self-Organizing Maps, *Tellus A*, 68, 29293, doi:10.3402/tellusa.v68.29293.
33. Pham X. T., Fontaine, B., and Philippon, N., 2010: Onset of the summer monsoon over southern Vietnam and its predictability. *Theor Appl Climatol*, 99, 105-113, <https://doi.org/10.1007/s00704-009-0115-z>
34. Pham-Thanh H., Phan-Van T., Fink A. H., and van der Linden R., 2021: Local-scale rainy season onset detection: A new approach based on principal component analysis and its application to Vietnam. *Int J Climatol*, 1–17. https://doi.org/10. 1002/joc.7441
35. Privé N. C., and Plumb R. A., 2007: Monsoon Dynamics with Interactive Forcing. Part I: Axisymmetric Studies, *J Atmos Sci*, *64*(5), <https://doi.org/10.1175/JAS3916.1>.
36. Slingo J., 2003: Overview. Monsoon. *Encyclopedia of Atmospheric Science*, Holton, J.R., J.A. Curry, J.A. Pyle, Academic Press, Vol.3, 1365-1369.
37. Thang N. V., Mau N. D., Khiem M. V., Duong T. H., Kham D. V., Thuy T. T., Tuan V. Q., Minh T. T. T., 2022: Climatic Factors Associated with Heavy Rainfall in Northern Vietnam in Boreal Spring, *Adv Meteorol*, vol. 2022, Article ID 5917729, 14 pages, <https://doi.org/10.1155/2022/5917729>
38. Thang N. V., Mau N. D., Van D. Q., Tuan B. M., Khiem M. V., Kham D. V., Thuy T. T., Duong T. H., Tam T. T., Quyen N. H., Thai L. X., Hien T. D., 2023: Orographic Effect and the Opposite Trend of Rainfall in Central Vietnam, *Adv Meteorol*, vol. 2023, Article ID 7256634, 12 pages, <https://doi.org/10.1155/2023/7256634>
39. Wallace J. M., & Hobbs P. V., 1977: Atmosphere science-an introductory survey. *Atmosphere science-an introductory survey*, V.
40. Wang B., LinHo, Zhang Y., and Lu M., 2004: Definition of South China Sea Monsoon Onset and Commencement of the East Asia Summer Monsoon. *J Climate*, 17, 699–710.
41. Webster P. J., Magaña V. O., Palmer T. N., Shukla J., Tomas R. A., Yanai M., and Yasunari T., 1998: Monsoons: Processes, predictability, and the prospects for prediction, *J Geophys Res Oceans*, 103( C7), 14451– 14510.
42. Zhang Y., Li T., Wang B., and Wu G., 2002: Onset of the Summer Monsoon over the Indochina Peninsula: Climatology and Interannual Variations. *J Climate*, 15, 3206–3221, [https://doi.org/10.1175/1520-0442(2002)015<3206:OOTSMO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015%3c3206:OOTSMO%3e2.0.CO;2).
43. Zhang Z., Chan J. C., and Ding Y., 2004: Characteristics, evolution and mechanisms of the summer monsoon onset over Southeast Asia. *Int J Climatol*, 24(12), 1461-1482, <https://doi.org/10.1002/joc.1082>.