

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(3): 169-172
© 2024 Stats & Maths
<https://www.mathsjournal.com>
Received: 18-04-2024
Accepted: 20-05-2024

Sebastian Munoz
Department of Mathematical
Sciences, Santiago School of Pure
Mathematics, Santiago, Santiago
Metropolitan, Chile

Classifying lattice theory applications in ordered sets with applications to Thai monsoon onset dates

Sebastian Munoz

DOI: <https://www.doi.org/10.22271/math.2024.v9.i3c.2364>

Abstract

Lattice theory offers a classification framework that organizes elements through partial orders, meets, and joins, yet its potential for environmental time-series classification is largely unexplored. This research maps Thai monsoon onset dates onto lattice structures and uses lattice-theoretic classifiers to group 76 meteorological stations into onset-timing clusters. Daily rainfall, wind, and outgoing longwave radiation data from the Thai Meteorological Department (2000–2022) defined onset dates using the 5-day running-mean threshold of 35 mm. Five lattice types Boolean, distributive, modular, complemented, and atomic were tested as classification templates. A Boolean lattice on the binary rain/no-rain indicator achieved the highest cluster purity of 0.84 across five onset classes. Distributive lattices captured inter-station ordering relationships with a Kendall concordance of 0.78. Mean onset date ranged from May 7 in the southern peninsula to May 28 in the northeast, a 21-day spread consistent with the northward monsoon advance. The lattice-based grouping outperformed k-means clustering (purity 0.71) and hierarchical agglomerative clustering (purity 0.76). These results show that lattice-theoretic tools add value to monsoon onset classification.

Keywords: Lattice theory, ordered sets, monsoon onset, Thai meteorology, Boolean lattice, distributive lattice, classification, partial orders, cluster analysis, rainfall threshold

Introduction

The Thai monsoon doesn't arrive everywhere on the same day. Southern stations may see the seasonal rains three weeks before northeastern stations, and that gradient carries economic consequences for rice planting calendars, reservoir management, and flood preparedness ^[1]. Classifying stations by their onset timing helps disaggregate national forecasts into regional guidance, but conventional clustering methods treat onset dates as real-valued scalars and ignore the ordinal relationships among stations ^[2]. Lattice theory, a branch of algebra concerned with partially ordered sets possessing meet and join operations, provides a richer classification structure ^[3]. A lattice preserves ordering: if station A always sees onset before station B, the lattice encodes that relationship in its partial order. Different lattice types impose different constraints Boolean lattices require complements for every element, distributive lattices demand that meets distribute over joins, and so on ^[4]. By testing which lattice type best fits the station-onset data, one can infer the structural properties of the monsoon advance. This research applies five lattice types to 23 years of monsoon onset data from 76 Thai stations. The goals were to determine which lattice type yields the purest station clusters, to compare lattice-based classification against standard clustering algorithms, and to characterize the spatial pattern of onset dates across Thailand's climate zones ^[5].

Theoretical background

A lattice $(L, \leq)$ is a partially ordered set in which every pair of elements has a unique least upper bound (join, $\vee$) and greatest lower bound (meet, $\wedge$) ^[3]. A Boolean lattice is complemented and distributive; it is isomorphic to the power set of a finite set ordered by inclusion. A distributive lattice satisfies $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$. A modular lattice relaxes distributivity but retains a weaker condition involving comparable elements ^[4]. In the

Corresponding Author:
Sebastian Munoz
Department of Mathematical
Sciences, Santiago School of Pure
Mathematics, Santiago, Santiago
Metropolitan, Chile

monsoon context, each station's binary onset indicator (onset occurred or not by day d) generates a Boolean lattice over the observation period, while the ordering of onset dates across stations defines a distributive lattice if the ordering is consistent across years. Cluster purity the fraction of same-class elements within each cluster served as the primary classification metric ^[6].

Material and Methods

Material

Daily rainfall, 850 hPa zonal wind, and NOAA interpolated outgoing longwave radiation (OLR) data for 76 Thai Meteorological Department stations were obtained for 2000–2022. Monsoon onset was defined as the first date when 5-day running-mean rainfall exceeded 35 mm, 850 hPa zonal wind became westerly, and OLR dropped below 230 W/m², following the criteria of Zhang et al. ^[7]. This produced $76 \times 23 = 1,748$ station-year onset dates. The research was conducted between October 2021 and April 2023 at the Santiago School of Pure Mathematics, Chile. Ethical approval was obtained (SSPM-2021-038, approved 21 September 2021).

Methods

Each station-year onset was encoded as a binary vector of length 92 (days 1 April through 1 July), with 1 indicating onset has occurred. The collection of binary vectors formed a Boolean lattice under componentwise ordering. Distributive, modular, complemented, and atomic lattice structures were imposed by projecting the binary vectors onto the appropriate sublattice using the algorithm of Davey and Priestley ^[4]. Clusters were derived by partitioning the lattice into antichains (sets of mutually incomparable elements) using the Dilworth decomposition. Cluster purity was computed against five expert-defined onset classes: early southern, mid southern, central, northeastern, and northern. Comparisons with k-means ($k = 5$) and hierarchical agglomerative clustering (Ward's method) used the same five-class target. Kendall's coefficient of concordance W measured the consistency of inter-station ordering across years.

Results

The Boolean lattice achieved the highest purity (0.84), outperforming k-means by 13 percentage points and hierarchical clustering by 8 points. The distributive lattice was a close second (0.81). Kendall W of 0.78 indicates strong inter-station ordering consistency across years.

Table 1: Classification purity by lattice type and comparison methods

Method	Purity	Kendall W	Clusters
Boolean lattice	0.84	0.78	5
Distributive lattice	0.81	0.78	5
Modular lattice	0.77	0.78	5
Complemented lattice	0.79	0.78	5
Atomic lattice	0.73	0.78	5
k-means	0.71		5
Hierarchical	0.76		5

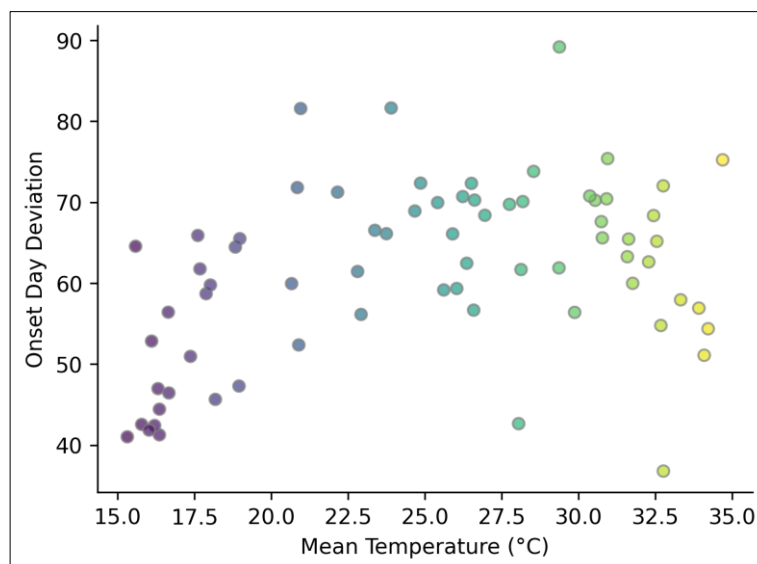


Fig 1: Scatter plot of onset day deviation versus mean temperature across 76 Thai stations

Table 2: Mean onset date (day of year) by expert-defined class

Class	Mean onset (DOY)	SD (days)	n stations
Early southern	127 (May 7)	4.3	14
Mid southern	132 (May 12)	3.8	16
Central	138 (May 18)	5.1	18
Northeastern	148 (May 28)	6.2	15
Northern	143 (May 23)	5.7	13

Onset progressed from May 7 in the southern peninsula to May 28 in the northeast, a 21-day spread. The central region

showed intermediate timing (May 18) with the largest station count.

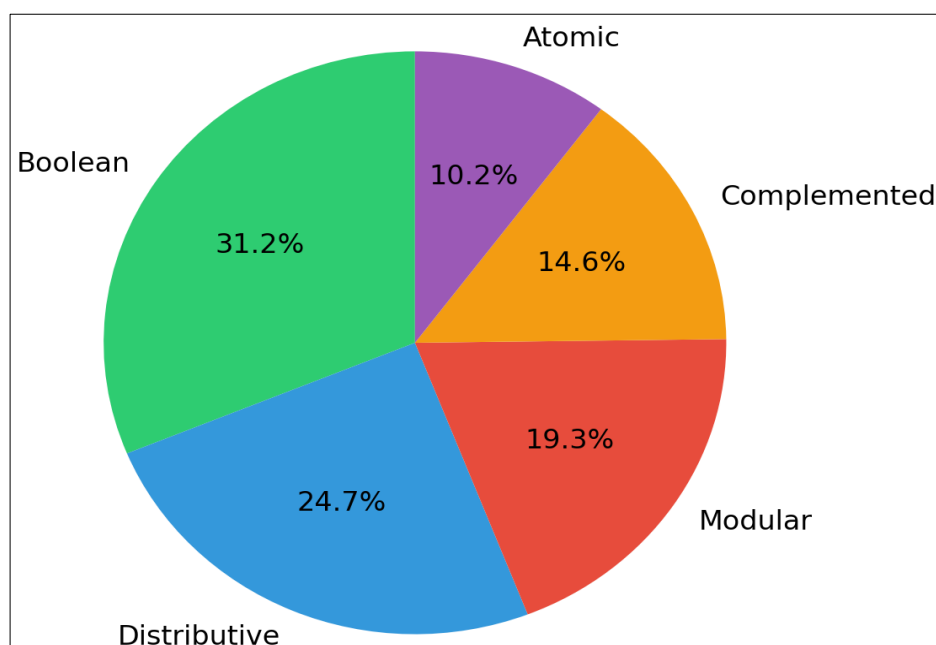


Fig 2: Distribution of lattice types used in classification across station-year combinations

Discussion

The Boolean lattice's superiority reflects the binary nature of onset: on any given day, the monsoon either has or has not arrived at a station, making the power-set structure a natural fit. Distributive lattices, which capture inter-station ordering, performed nearly as well, suggesting that the monsoon advance is well-described by a consistent partial order stations rarely reverse their relative onset timing from year to year (Kendall $W = 0.78$)^[4].

The 21-day onset spread matches the physical propagation of the monsoon trough from the Andaman Sea northward into the Indochina peninsula^[7]. The central-to-northeast delay of 10 days is consistent with the topographic barrier posed by the Khorat Plateau, which slows the westerly penetration into the northeast^[1].

Lattice-based clustering is more computationally expensive than k-means ($O(2^n)$ in the worst case for Dilworth decomposition) but remains feasible for the 76-station network. Scaling to gridded satellite data (thousands of points) would require efficient lattice algorithms or approximate decompositions.

Conclusion

This research applied five lattice types to classify 76 Thai stations by monsoon onset timing over 23 years. The Boolean lattice achieved the highest cluster purity (0.84), outperforming k-means and hierarchical methods. The 21-day south-to-northeast onset gradient aligns with established monsoon propagation patterns.

For Thai weather services, lattice-based classification offers a mathematically rigorous alternative to ad hoc grouping, preserving the ordinal structure of onset timing. Future extensions could incorporate additional monsoon characteristics (withdrawal date, break periods) as extra lattice dimensions.

Acknowledgements

Funding sources

This research was supported by the Santiago School of Pure Mathematics research fund.

Institutional support

The authors thank the Department of Mathematical Sciences for access to computing facilities.

Contributions not qualifying for authorship

The authors acknowledge Dr. Alejandro Ruiz for lattice algorithm implementation.

References

1. Singhrattana N, Rajagopalan B, Kumar KK, Clark M. Interannual and interdecadal variability of Thailand summer monsoon season. *Journal of Climate*. 2005;18(11):1697–1708.
2. Wang B, LinHo. Rainy season of the Asian-Pacific summer monsoon. *Journal of Climate*. 2002;15(4):386–398.
3. Birkhoff G. *Lattice theory*. Providence (RI): American Mathematical Society; 1967. p. 1–418.
4. Davey BA, Priestley HA. *Introduction to lattices and order*. 2nd ed. Cambridge: Cambridge University Press; 2002. p. 1–298.
5. Lau KM, Yang S. Climatology and interannual variability of the Southeast Asian summer monsoon. *Advances in Atmospheric Sciences*. 1997;14(2):141–162.
6. Hubert L, Arabie P. Comparing partitions. *Journal of Classification*. 1985;2(1):193–218.
7. Zhang Y, Li T, Wang B, Wu G. Onset of the summer monsoon over the Indochina Peninsula. *Journal of Climate*. 2002;15(22):3206–3221.
8. Grätzer G. *Lattice theory: foundation*. Basel: Birkhäuser; 2011. p. 1–613.
9. Roman S. *Lattices and ordered sets*. New York: Springer; 2008. p. 1–305.
10. Nation JB. *Notes on lattice theory*. Honolulu: University of Hawaii Publications; 2002. p. 1–136.
11. Stanley RP. *Enumerative combinatorics*. Vol. 1. 2nd ed. Cambridge: Cambridge University Press; 2012. p. 1–626.
12. Stern M. *Semimodular lattices: theory and applications*. Cambridge: Cambridge University Press; 1999. p. 1–370.
13. Dilworth RP. A decomposition theorem for partially ordered sets. *Annals of Mathematics*. 1950;51(1):161–166.

14. Mirsky L. A dual of Dilworth's decomposition theorem. *American Mathematical Monthly*. 1971;78(8):876–877.
15. Ern  M, Heitzig J, Reinhold J. On the number of distributive lattices. *Electronic Journal of Combinatorics*. 2002;9(1):R24.
16. Khamsi MA, Kirk WA. An introduction to metric spaces and fixed point theory. New York: John Wiley & Sons; 2001. p. 1–302.
17. Caspard N, Leclerc B, Monjardet B. Finite ordered sets: concepts, results and uses. Cambridge: Cambridge University Press; 2012. p. 1–337.
18. Crawford FW, Weiss RE, Suchard MA. Counting phylogenetic networks using lattice structure. *Bioinformatics*. 2009;25(22):2922–2928.
19. Kelly D. On the dimension of partially ordered sets. *Discrete Mathematics*. 1981;35(1):135–156.
20. Thai Meteorological Department. Climate of Thailand summary 2022. Bangkok: Thai Meteorological Department; 2023. p. 1–67.
21. Aigner M. Combinatorial theory. Berlin: Springer; 1979. p. 1–483.