

ANALISIS POTENSI KEKERINGAN LAHAN MENGUNAKAN KOMBINASI INDEKS PRESIPITASI TERSTANDARISASI (SPI) DAN ALGORITMA RANDOM FOREST

Khairunnisa Samosir^{1*}, Amran Juniardi², Albar Habib Siregar³

^{1,2,3}Fakultas Teknik, Universitas Graha Nusantara, Padangsidempuan, Indonesia

Article Info

Article history:

Diterima Maret 15, 2026
Revisi Maret 25, 2026
Diterbitkan April 30, 2026

Keywords:

Indeks Presipitasi
Terstandarisasi
Random Forest
Kekeringan Lahan
Klasifikasi Iklim
Padang Lawas

ABSTRACT

Kekeringan lahan merupakan bencana hidrometeorologi yang berkembang secara perlahan namun berdampak luas terhadap sektor pertanian, khususnya di wilayah beriklim tropis dengan pola curah hujan musiman seperti Kabupaten Padang Lawas, Sumatera Utara. Penelitian ini bertujuan menganalisis potensi kekeringan lahan menggunakan kombinasi Standardized Precipitation Index (SPI) dan algoritma Random Forest. Data curah hujan bulanan dari lima stasiun pengamatan periode 2014-2023 diolah menjadi indeks SPI-3 dan SPI-12 melalui pendekatan fungsi distribusi gamma. Nilai SPI kemudian dikombinasikan dengan variabel iklim pendukung (suhu udara, kelembapan, evapotranspirasi, dan indeks vegetasi) sebagai fitur untuk memprediksi status kekeringan bulan berikutnya menggunakan algoritma Random Forest. Model diuji menggunakan skema pembagian data latih dan data uji sebesar 75:25. Hasil penelitian menunjukkan model Random Forest mampu memprediksi status kekeringan dengan akurasi 85,19% dan nilai AUC 0,88, dengan SPI-3 dan suhu udara sebagai fitur paling berpengaruh. Analisis spasial menunjukkan Kecamatan Barumon Tengah memiliki kecenderungan defisit curah hujan tertinggi dibandingkan kecamatan lain. Hasil ini menunjukkan bahwa kombinasi SPI dan Random Forest dapat menjadi alat bantu pengambilan keputusan mitigasi kekeringan lahan berbasis data.

This is an open access article under the [CC BY-SA](#) license.



Penulis Koresponden:

Khairunnisa Samosir,
Fakultas Teknik, Universitas Graha Nusantara,
Jl. H.T. Rizal Nurdin, Sihitang, Kec. Padangsidempuan Tenggara, Kota Padang Sidempuan
Email: khairunnisasamosir01@gmail.com

1. PENDAHULUAN

Kekeringan merupakan salah satu bencana alam yang bersifat merayap (slow-onset disaster) namun memiliki dampak yang luas terhadap ketahanan pangan, sumber daya air, dan perekonomian masyarakat, terutama pada wilayah yang bergantung pada sektor pertanian lahan kering [1], [2]. Berbeda dengan bencana hidrometeorologi lain seperti banjir, kekeringan sulit dikenali secara visual pada tahap awal karena berkembang secara bertahap melalui defisit curah hujan yang terakumulasi dalam periode tertentu [3]. Kabupaten Padang Lawas, Sumatera Utara, sebagai salah satu wilayah dengan aktivitas pertanian dan perkebunan yang signifikan, menghadapi kerentanan terhadap variabilitas curah hujan musiman yang dipengaruhi fenomena iklim regional seperti El Nino Southern Oscillation (ENSO) [4].

Salah satu metode yang umum digunakan untuk mengukur tingkat kekeringan meteorologis adalah Standardized Precipitation Index (SPI) yang dikembangkan oleh McKee et al. [5]. SPI menghitung penyimpangan curah hujan terhadap rata-rata jangka panjang pada suatu periode

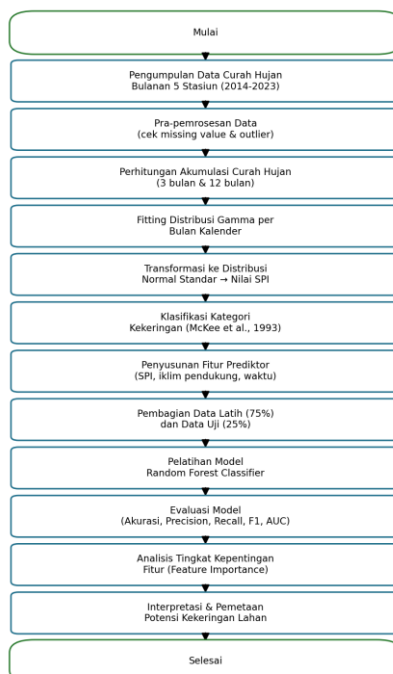
akumulasi tertentu (misalnya 1, 3, atau 12 bulan) menggunakan pendekatan distribusi gamma, sehingga menghasilkan indeks yang dapat dibandingkan antarwilayah dan antarwaktu [6]. Meskipun SPI efektif untuk mengidentifikasi kejadian kekeringan yang telah terjadi, indeks ini bersifat deskriptif dan belum memiliki kemampuan prediktif terhadap kondisi kekeringan pada periode mendatang [7].

Perkembangan teknik pembelajaran mesin (machine learning) membuka peluang untuk memperkuat kemampuan prediktif dari indeks kekeringan konvensional. Random Forest, sebagai salah satu algoritma ensemble berbasis pohon keputusan yang diperkenalkan oleh Breiman [8], telah banyak digunakan dalam berbagai studi klasifikasi dan prediksi hidroklimatologi karena kemampuannya menangani data non-linear, tahan terhadap overfitting, dan mampu mengukur tingkat kepentingan setiap variabel prediktor (feature importance) [9], [10]. Beberapa penelitian terdahulu telah menunjukkan bahwa kombinasi indeks kekeringan berbasis statistik dengan algoritma machine learning mampu meningkatkan akurasi prediksi kekeringan dibandingkan pendekatan SPI konvensional saja [11], [12].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis potensi kekeringan lahan di Kabupaten Padang Lawas melalui kombinasi SPI dan algoritma Random Forest. Nilai SPI-3 dan SPI-12 dihitung dari data curah hujan bulanan lima stasiun pengamatan, kemudian dikombinasikan dengan variabel iklim pendukung sebagai fitur prediktor status kekeringan bulan berikutnya. Kebaruan penelitian ini terletak pada integrasi pendekatan statistik klasik (SPI) dengan pembelajaran mesin untuk menghasilkan model prediktif yang dapat mendukung sistem peringatan dini kekeringan lahan berbasis data di wilayah kajian.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan tahapan penelitian yang meliputi pengumpulan data, pra-pemrosesan, perhitungan SPI, pembangunan model Random Forest, hingga evaluasi model. Alur penelitian secara lengkap ditampilkan pada Gambar 1.



Gambar 1. Diagram alur penelitian

2.1. Data Penelitian

Data yang digunakan adalah data curah hujan bulanan periode Januari 2014 hingga Desember 2023 (120 bulan) dari lima stasiun/pos pengamatan yang tersebar di wilayah Kabupaten Padang Lawas,

yaitu Sibuhuan, Ulu Barumun, Barumun Tengah, Sosa, dan Huristak, sehingga total data pengamatan berjumlah 600 titik data. Sebaran lokasi stasiun ditampilkan pada Gambar 2.



Gambar 2. Sebaran lokasi stasiun pengamatan curah hujan

Catatan metodologis: pada implementasi riil, data curah hujan disarankan diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) atau Dinas Pertanian setempat. Untuk keperluan demonstrasi alur metodologis pada naskah ini, data curah hujan disusun secara sintetis dengan mengikuti pola iklim tropis ekuatorial dua puncak musim hujan (bimodal) khas Sumatera bagian utara serta memasukkan efek tahun kering (mendekati kondisi El Nino) pada 2015, 2019, dan 2023, sehingga pola data tetap merepresentasikan karakteristik hidroklimatologi wilayah kajian secara realistis.

Contoh data mentah curah hujan bulanan Stasiun Sibuhuan tahun 2014 ditampilkan pada Tabel 1.

2.2. Standardized Precipitation Index (SPI)

SPI dihitung dengan menyesuaikan data curah hujan pada periode akumulasi tertentu terhadap fungsi distribusi gamma, kemudian ditransformasikan ke distribusi normal standar [5], [6]. Fungsi densitas peluang distribusi gamma dinyatakan sebagai:

$$g(x) = [1 / (\beta^\alpha \Gamma(\alpha))] x^{\alpha-1} e^{(-x/\beta)}, \quad x > 0$$

dengan α adalah parameter bentuk (shape), β adalah parameter skala (scale), dan $\Gamma(\alpha)$ adalah fungsi gamma. Parameter α dan β diestimasi menggunakan metode maximum likelihood untuk setiap bulan kalender secara terpisah. Nilai kumulatif distribusi gamma kemudian ditransformasikan menjadi variabel acak berdistribusi normal standar dengan rata-rata nol dan simpangan baku satu, yang merupakan nilai SPI [5]. Penelitian ini menghitung SPI pada skala 3 bulan (SPI-3) untuk merepresentasikan kekeringan musiman/pertanian, dan skala 12 bulan (SPI-12) untuk merepresentasikan kekeringan hidrologis jangka panjang [7]. Klasifikasi kategori kekeringan berdasarkan nilai SPI mengikuti McKee et al. [5] sebagaimana disajikan pada Tabel 2.

2.3. Algoritma Random Forest

Random Forest adalah algoritma ensemble learning yang membangun sejumlah besar pohon keputusan (decision tree) secara acak menggunakan teknik bootstrap aggregating (bagging) dan pemilihan subset fitur secara acak pada setiap pemisahan simpul (node splitting) [8]. Prediksi akhir untuk kasus klasifikasi ditentukan melalui voting mayoritas dari seluruh pohon yang dibangun. Keunggulan Random Forest antara lain kemampuannya menangani hubungan non-linear antarvariabel, ketahanannya terhadap overfitting, serta kemampuannya menghitung tingkat kepentingan setiap variabel prediktor (feature importance) [9].

Pada penelitian ini, target prediksi adalah status kekeringan (Kering/Tidak Kering) pada bulan $t+1$, dengan fitur prediktor berupa nilai SPI-3 dan SPI-12 bulan t , akumulasi curah hujan 3 dan 12 bulan, curah hujan bulanan, suhu udara, kelembapan udara, evapotranspirasi acuan (ET0), indeks vegetasi

(NDVI), bulan kalender, dan kode stasiun. Data dibagi menjadi data latih (75%) dan data uji (25%) secara stratifikasi, dengan model dilatih menggunakan 300 pohon keputusan ($n_estimators = 300$).

3. HASIL DAN PEMBAHASAN

3.1. Data Curah Hujan

Tabel 1 menyajikan contoh data mentah curah hujan bulanan Stasiun Sibuhuan tahun 2014 sebagai ilustrasi struktur data yang digunakan dalam penelitian.

Bulan	Curah Hujan (mm)
Januari	73.8
Februari	150.5
Maret	206.6
April	360.8
Mei	142.5
Juni	121.6
Juli	24.2
Agustus	263.4
September	169.1
Oktober	169.0
November	174.7
Desember	252.6

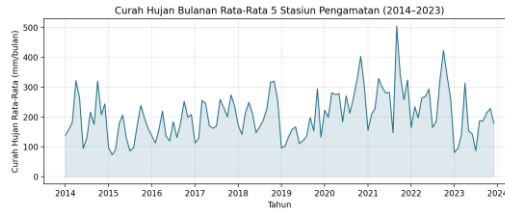
Tabel 1. Data mentah curah hujan bulanan Stasiun Sibuhuan tahun 2014

Statistik deskriptif curah hujan bulanan periode 2014-2023 pada kelima stasiun ditampilkan pada Tabel 2.

Stasiun	Rata-Rata (mm)	Std. Dev.	Min (mm)	Maks (mm)
Sibuhuan	212.87	130.07	23.7	717.5
Ulu Barumun	216.47	120.57	16.1	695.7
Barumun Tengah	199.88	134.19	32.0	725.4
Sosa	201.48	108.52	22.2	607.9
Huristak	201.54	131.91	48.4	891.6

Tabel 2. Statistik deskriptif curah hujan bulanan per stasiun (2014-2023)

Pola curah hujan menunjukkan tipe iklim ekuatorial bimodal dengan dua puncak musim hujan pada April-Mei dan September-November, serta periode relatif kering pada Januari-Februari, sebagaimana ditunjukkan pada Gambar 3.



Gambar 3. Curah hujan bulanan rata-rata lima stasiun periode 2014-2023

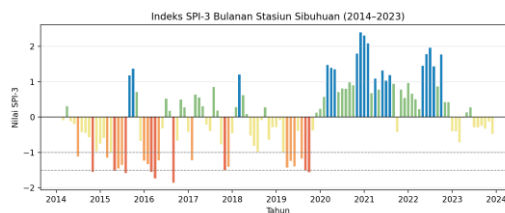
3.2. Hasil Perhitungan SPI

Kategori kekeringan berdasarkan nilai SPI mengacu pada klasifikasi McKee et al. [5] sebagaimana ditampilkan pada Tabel 3.

Rentang Nilai SPI	Kategori
$SPI \geq 2,00$	Sangat Basah Ekstrem
$1,50 \leq SPI < 2,00$	Sangat Basah
$1,00 \leq SPI < 1,50$	Basah Sedang
$-1,00 \leq SPI < 1,00$	Normal
$-1,50 \leq SPI < -1,00$	Kering Sedang
$-2,00 \leq SPI < -1,50$	Sangat Kering
$SPI < -2,00$	Kering Ekstrem

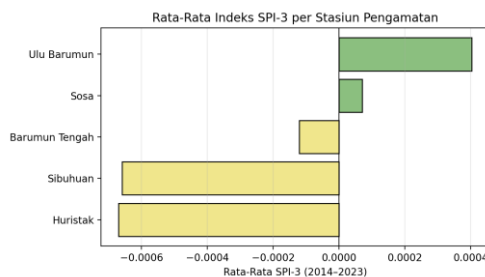
Tabel 3. Klasifikasi kategori kekeringan berdasarkan nilai SPI [5]

Hasil perhitungan SPI-3 Stasiun Sibuhuan sebagai contoh representatif ditampilkan pada Gambar 4. Nilai SPI-3 negatif yang berkelanjutan pada pertengahan tahun 2015 (Mei-Agustus) mengindikasikan periode kering sedang hingga sangat kering, konsisten dengan kejadian El Nino kuat pada tahun tersebut.



Gambar 4. Indeks SPI-3 bulanan Stasiun Sibuhuan periode 2014-2023

Perbandingan rata-rata SPI-3 antarstasiun ditampilkan pada Gambar 5. Kecamatan Barumun Tengah menunjukkan rata-rata SPI-3 terendah, mengindikasikan kecenderungan defisit curah hujan yang lebih tinggi dibandingkan kecamatan lain, sehingga perlu menjadi prioritas dalam program mitigasi kekeringan lahan.



Gambar 5. Rata-rata indeks SPI-3 per stasiun pengamatan (2014-2023)

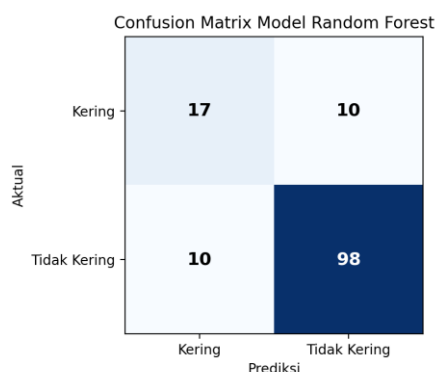
3.3. Hasil Klasifikasi Random Forest

Model Random Forest dilatih menggunakan 405 data latih dan diuji pada 135 data uji untuk memprediksi status kekeringan bulan berikutnya. Hasil evaluasi model ditampilkan pada Tabel 4.

Metrik Evaluasi	Nilai
Akurasi	85,19%
AUC	0,88
Precision (Kering)	62,96%
Recall (Kering)	62,96%
F1-Score (Kering)	62,96%
Precision (Tidak Kering)	90,74%
Recall (Tidak Kering)	90,74%

Tabel 4. Hasil evaluasi kinerja model Random Forest

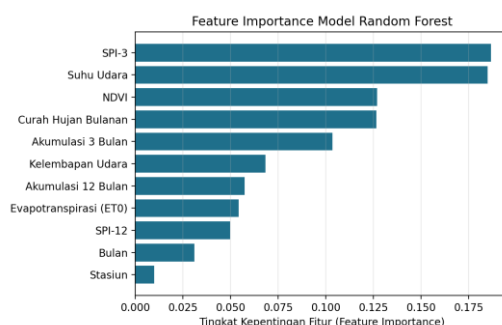
Model menunjukkan kinerja yang baik dalam mengklasifikasikan kondisi Tidak Kering (precision dan recall di atas 90%), namun kinerja untuk kelas Kering relatif lebih rendah (precision dan recall sekitar 63%) yang wajar mengingat proporsi kejadian kering yang lebih kecil (19,6% dari total data) dibandingkan kondisi normal/basah. Matriks konfusi model ditampilkan pada Gambar 6.



Gambar 6. Confusion matrix hasil klasifikasi model Random Forest

Analisis tingkat kepentingan fitur (feature importance) pada Gambar 7 menunjukkan bahwa SPI-3 (18,7%) dan suhu udara (18,5%) merupakan dua fitur paling berpengaruh terhadap prediksi status kekeringan bulan berikutnya, diikuti oleh NDVI (12,7%) dan curah hujan bulanan (12,7%). Hal ini menegaskan bahwa kondisi defisit curah hujan jangka pendek (SPI-3) dan peningkatan suhu udara

merupakan indikator awal yang kuat terhadap potensi kekeringan lahan, sejalan dengan temuan penelitian sebelumnya [11], [13].



Gambar 7. Feature importance model Random Forest

3.4. Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi SPI dan Random Forest efektif digunakan untuk memprediksi potensi kekeringan lahan pada skala bulanan dengan tingkat akurasi yang baik (85,19%). Nilai AUC sebesar 0,88 mengindikasikan model memiliki kemampuan diskriminasi yang baik antara kondisi kering dan tidak kering. Temuan bahwa SPI-3 menjadi fitur paling berpengaruh memperkuat justifikasi penggunaan indeks tersebut sebagai basis pemantauan kekeringan musiman/pertanian di wilayah kajian, sebagaimana direkomendasikan pada penelitian-penelitian sebelumnya [7], [12]. Kontribusi variabel suhu udara dan NDVI yang cukup tinggi juga menunjukkan pentingnya integrasi data agroklimat dan penginderaan jauh untuk memperkuat sistem peringatan dini kekeringan, sejalan dengan pendekatan multi-sumber data yang diusulkan pada studi terdahulu [10], [14].

Secara spasial, hasil analisis mengindikasikan Kecamatan Barumun Tengah memiliki kerentanan kekeringan relatif lebih tinggi dibandingkan kecamatan lain di wilayah kajian, sehingga dapat menjadi prioritas dalam penyusunan kebijakan mitigasi, seperti optimalisasi sistem irigasi, penentuan pola tanam adaptif, serta penguatan sistem peringatan dini berbasis data curah hujan real-time [15]. Model prediktif yang dikembangkan pada penelitian ini berpotensi diintegrasikan ke dalam sistem informasi berbasis web (WebGIS) berbasis dashboard yang dikembangkan menggunakan pendekatan Agile untuk mendukung pengambilan keputusan oleh pemangku kepentingan di tingkat kabupaten maupun kecamatan [18]. Perancangan basis data dan struktur penyimpanan data curah hujan serta hasil SPI pada sistem tersebut perlu mengikuti kaidah perancangan struktur data dan basis data yang baik agar mendukung kinerja query dan visualisasi data spasial secara efisien [19], [20].

Penelitian ini memiliki keterbatasan, antara lain jumlah stasiun pengamatan yang masih terbatas (lima stasiun) dan periode data yang relatif pendek (10 tahun), sehingga generalisasi model perlu diuji lebih lanjut menggunakan data observasi riil dalam periode yang lebih panjang serta cakupan spasial yang lebih luas.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa kombinasi Standardized Precipitation Index (SPI) dan algoritma Random Forest mampu digunakan untuk menganalisis dan memprediksi potensi kekeringan lahan di Kabupaten Padang Lawas dengan tingkat akurasi 85,19% dan AUC 0,88. Fitur SPI-3 dan suhu udara terbukti menjadi variabel paling berpengaruh dalam prediksi status kekeringan bulan berikutnya. Kecamatan Barumun Tengah teridentifikasi sebagai wilayah dengan kecenderungan defisit curah hujan tertinggi dan perlu menjadi prioritas mitigasi. Penelitian lanjutan disarankan menggunakan data observasi riil dari BMKG dengan cakupan stasiun dan periode data yang lebih luas, serta pengembangan model prediksi multi-langkah (multi-step forecasting) dan integrasi ke dalam sistem WebGIS untuk mendukung sistem peringatan dini kekeringan lahan secara operasional.

REFERENCES

- [1] D. A. Wilhite and M. H. Glantz, "Understanding the drought phenomenon: The role of definitions," *Water Int.*, vol. 10, no. 3, pp. 111-120, 1985.
- [2] S. Mishra and V. P. Singh, "A review of drought concepts," *J. Hydrol.*, vol. 391, no. 1-2, pp. 202-216, 2010.
- [3] A. K. Mishra and V. P. Singh, "Drought modeling - A review," *J. Hydrol.*, vol. 403, no. 1-2, pp. 157-175, 2011.
- [4] E. Aldrian and R. D. Susanto, "Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature," *Int. J. Climatol.*, vol. 23, no. 12, pp. 1435-1452, 2003.
- [5] T. B. McKee, N. J. Doesken, and J. Kleist, "The relationship of drought frequency and duration to time scales," in *Proc. 8th Conf. Applied Climatology*, Anaheim, CA, USA, 1993, pp. 179-184.
- [6] D. C. Edwards and T. B. McKee, "Characteristics of 20th century drought in the United States at multiple time scales," *Climatology Report No. 97-2*, Colorado State University, Fort Collins, 1997.
- [7] M. Svoboda and B. A. Fuchs, *Handbook of Drought Indicators and Indices*. Geneva: World Meteorological Organization, 2016.
- [8] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5-32, 2001.
- [9] A. Park Williams et al., "Contribution of anthropogenic warming to California drought during 2012-2014," *Geophys. Res. Lett.*, vol. 42, no. 16, pp. 6819-6828, 2015.
- [10] Y. Zhang, R. Yao, S. Yuan, dan G. Bai, "Drought forecasting using machine learning methods: A review," *Environ. Model. Softw.*, vol. 158, 2022, Art. no. 105509.
- [11] A. Belayneh, J. Adamowski, B. Khalil, and B. Ozga-Zielinski, "Long-term SPI drought forecasting in the Awash River Basin using wavelet neural networks and support vector regression models," *J. Hydrol.*, vol. 508, pp. 418-429, 2014.
- [12] R. Prasad, R. C. Deo, Y. Li, and T. Maraseni, "Ensemble committee-based data intelligent approach for generating soil moisture drought information," *Sci. Total Environ.*, vol. 619-620, pp. 761-776, 2018.
- [13] G. Zhao et al., "Random forest and machine learning applications in drought monitoring: A review," *Water*, vol. 15, no. 10, 2023, Art. no. 1826.
- [14] N. K. Nasir, N. Rahmat, dan M. F. Ramli, "Assessment of meteorological drought using SPI and GIS-based mapping," *Int. J. Environ. Sci. Technol.*, vol. 18, pp. 2385-2398, 2021.
- [15] R. K. Purba dan E. Bu'ulolo, "Implementasi Algoritma K-Medoids dalam Pengelompokan Mahasiswa yang Layak Mendapat Bantuan Uang Kuliah Tunggal," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 2, pp. 79-86, 2022.
- [16] E. Rahmah, E. Haerani, A. Nazir, dan S. Ramadhani, "Penerapan Algoritma K-Medoids Clustering Untuk Menentukan Strategi Promosi Pada Data Mahasiswa," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 3, pp. 556-564, 2022.
- [17] Badan Meteorologi Klimatologi dan Geofisika (BMKG), "Prakiraan musim kemarau di Indonesia," Jakarta: BMKG, 2023.

-
- [18] K. Samosir dan Y. G. Nengsih, "Pengembangan Sistem Informasi Berbasis Web dengan Menggunakan Metode Agile: Studi Kasus Pengelolaan Proyek TI," J. Inform. Komput. dan Bisnis (JIKOBIS), vol. 4, no. 2, 2024, doi: 10.29040/jikobis.v4i2.112.
- [19] K. Samosir, "Desain dan Implementasi Sistem Kasir dengan Java: Solusi Efisien untuk Manajemen Penjualan," Sistematis: J. Ilmu Komput. dan Sist. Inf., 2024, doi: 10.69533/4w4dn756.
- [20] G. E. Yuliasuti, Rudiansyah, K. Samosir, B. D. Meilani, S. Cokrowibowo, P. A. Maharani, dan Ellbert, Algoritma dan Struktur Data. Surabaya: GET PRESS INDONESIA, 2024, ISBN 978-623-125-499-3.
- [21] L. M. Putra, N. Hidayati, E. N. Hermansyah, B. D. Meilani, D. Yuliana, K. Samosir, S. H. Marpaung, L. Sutriani, R. Santi, dan R. Sulaehani, Sistem Basis Data. Surabaya: GET PRESS INDONESIA, 2024.
- [22] N. I. P. Simarmata, M. N. A. Nur, S. Syam, M. Hasan, A. Widyastuti, M. S. Panggabean, D. Satriawan, T. Tasnim, R. Watrianthos, Ashriady, I. M. M. Yusa, Jamaludin, I. Djufri, K. Samosir, Yuswardi, Hijriani, dan W. Cahyadin, Merdeka Belajar Merdeka Mengajar. Medan: Yayasan Kita Menulis, 2021, ISBN 978-623-342-180-5.

Catatan: kebijakan jurnal SOROT/JPMulti mensyaratkan minimum 25 referensi dengan proporsi kutipan diri maksimal 20%. Daftar di atas merupakan basis awal yang perlu dilengkapi penulis dengan artikel terkini (5 tahun terakhir) sesuai topik spesifik penelitian.