

Estimasi Konsentrasi PM₁₀ Menggunakan *Support Vector Regression*

PM₁₀ Concentration Estimation Using Support Vector Regression

Zia Ayu Frakusya¹, Rista Hernandi Virgianto^{2*}, Muhammad Elifant Yuggotomo³

^{1,2}Sekolah Tinggi Meteorologi Klimatologi dan Geofisika
Jl. Perhubungan I No. 5, Pondok Betung 15221, Indonesia

³Badan Meteorologi Klimatologi dan Geofisika

*email: rista.virgianto@stmkg.ac.id

ABSTRAK

DOI:
10.30595/jrst.v6i1.8977

Histori Artikel:

Diajukan:
06/11/2020

Diterima:
03/10/2022

Diterbitkan:
11/11/2022

PM₁₀ berkontribusi terhadap polusi udara pada saat kejadian kabut asap di musim kemarau dengan salah satu sumber utamanya adalah pembakaran biomassa. Pada saat musim kemarau, terdapat banyak kegiatan pembersihan lahan di Mempawah untuk persiapan masa tanam yang sebagian besar dilakukan dengan pembakaran sisa tanaman. Sebagai salah satu polutan utama yang dapat memengaruhi kesehatan manusia, maka estimasi konsentrasi PM₁₀ sangat penting untuk dilakukan. Tujuan dari penelitian ini untuk melakukan estimasi konsentrasi PM₁₀ di Mempawah tahun 2019 menggunakan *Support Vector Regression* (SVR) berdasarkan data PM₁₀ dan variabel meteorologi seperti curah hujan, kelembaban, suhu, tekanan permukaan laut dan kecepatan angin dari Stasiun Klimatologi Mempawah dengan periode latih tahun 2016 hingga 2018. Hasil penelitian ini menunjukkan bahwa estimasi konsentrasi PM₁₀ menggunakan SVR berdasarkan parameter cuaca dapat menggambarkan variabilitas konsentrasi harian PM₁₀ di Mempawah dengan baik, terkecuali saat terjadi kenaikan konsentrasi yang sangat tinggi yang mungkin dipengaruhi oleh faktor antropogenik. Selain itu, berdasarkan verifikasi, RMSE yang dihasilkan model estimasi hampir sama dengan nilai standar deviasi observasinya.

Kata Kunci: PM₁₀, *Support Vector Regression*, Kualitas Udara, Estimasi, Kabut Asap

ABSTRACT

PM₁₀ contributes to air pollution during haze events in the dry season with one of the main sources being biomass burning. During the dry season, there are many land clearing activities in Mempawah to prepare for planting, mostly by burning crop residues. As one of the main pollutants that can affect human health, it is important to estimate the PM₁₀ concentration. The purpose of this study is to estimate the concentration of PM₁₀ in Mempawah in 2019 using *Support Vector Regression* (SVR) based on PM₁₀ data and meteorological variables such as rainfall, humidity, temperature, sea level pressure and wind speed from the Mempawah Climatology Station with a training period from 2016 to 2018. The results of this study indicate that the estimation of PM₁₀ concentrations using SVR based on weather parameters can describe the variability of PM₁₀ daily concentrations in Mempawah well, except when there is a very high increase in concentration which may be influenced by anthropogenic factors. In addition, based on verification, the RMSE generated by the estimation model is almost the same as the standard deviation value of the observations.

Keywords: PM₁₀, *Support Vector Regression*, Air Quality, Estimation, Smog

1. PENDAHULUAN

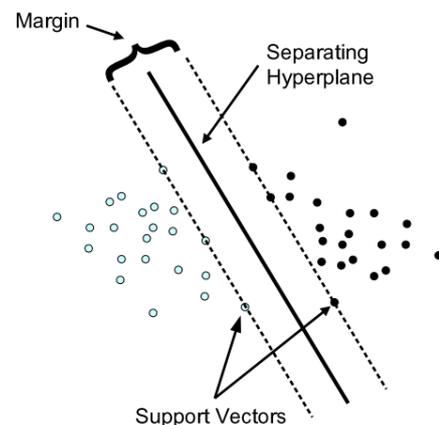
Paparan tingkat polusi udara yang tinggi memiliki dampak langsung pada manusia (W.-Z. Lu & Wang, 2005; W. Lu et al., 2002). Tingginya tingkat polusi udara di atmosfer sangat dipengaruhi oleh parameter iklim dan perilaku polutan (W. Lu et al., 2002). Salah satu bagian dari polusi udara adalah materi partikulat (PM) yang merupakan campuran partikel padat, cair atau padat dan cair yang tersuspensi di udara (WHO, 2003). Potensi materi partikulat dalam memengaruhi kesehatan manusia dapat dikaitkan dengan ukuran partikel, salah satunya adalah PM₁₀ atau materi partikulat yang berukuran 10 µm atau lebih kecil (EPA, 2019). PM₁₀ merupakan salah satu polutan utama yang dapat memengaruhi kesehatan manusia (Martin, 2011; Ul-Saufie, Yahaya, Ramli, Rosaida, & Hamid, 2013). Partikel ini umumnya masuk melewati tenggorokan dan hidung kemudian paru-paru. Apabila partikel tersebut terhirup maka dapat memengaruhi kesehatan yang serius (EPA, 2019).

Polusi udara di Kalimantan dan Sumatera disebabkan oleh kabut asap akibat pembakaran lahan gambut (Tacconi, 2003). Penyebab kebakaran paling umum dalam pembukaan lahan adalah penggunaan metode tebang dan bakar (Yulianti, 2018). Namun, sayangnya masyarakat di Kabupaten Mempawah, Provinsi Kalimantan Barat pada musim kemarau memiliki kebiasaan membuka lahan dengan cara membakar untuk persiapan masa tanam (Meilantika, Hadisaputro, & Setiawati, 2018). Menurut Badan Penanggulangan Bencana Daerah (BPBD) Mempawah, kebakaran hutan dan lahan di Kabupaten Mempawah per 9 Agustus 2019 telah mencapai 473 hektar (RRI.co.id., 2019). Selain itu, Dinas Kesehatan Kabupaten Mempawah menyampaikan bahwa kondisi udara di Kabupaten Mempawah terkait kabut asap dalam keadaan tidak sehat berdasarkan alat pengamat kualitas udara PM₁₀ di Stasiun Klimatologi Mempawah. Hal tersebut menyebabkan kasus penderita ISPA semakin meningkat sekitar 681 kasus (Suarapemredkalbar.com., 2019).

PM₁₀ berdasarkan penelitian Chantara (2012) di Thailand Utara dianggap sebagai polutan udara yang paling signifikan berkontribusi dalam polusi udara dengan salah satu sumber utamanya adalah pembakaran terbuka di musim kemarau dan bertepatan pula dengan puncak tahunan kabut asap di wilayah atas Thailand Utara. Selain itu, analisis konsentrasi PM₁₀ cocok untuk dilakukan pada

saat kejadian kabut asap akibat dari kebakaran hutan dan lahan (Munir, Akbar, Badaruddin, & Wahdah, 2018).

Memprediksi polusi udara untuk saat ini selain dengan menggunakan metode konvensional seperti metode statistik dan Jaringan Saraf Tiruan (JST) dapat pula menggunakan *Support Vector Machine* (SVM) (W. Wang, Men, & Lu, 2008). SVM yang ditunjukkan pada Gambar 1 diusulkan oleh Vapnik yang merupakan jenis klasifikasi berdasarkan teknik pembelajaran statistik (Vapnik, 2013). Tujuan dari SVM adalah memisahkan kelas dengan *hyperplane* yang memaksimalkan margin antara kelas tersebut (X. Wang & Zhong, 2003). SVM juga sudah dikembangkan untuk mengatasi masalah data non-linear dengan menggunakan *kernel* pada ruang berdimensi tinggi (Nugroho, Witarto, & Handoko, 2003).

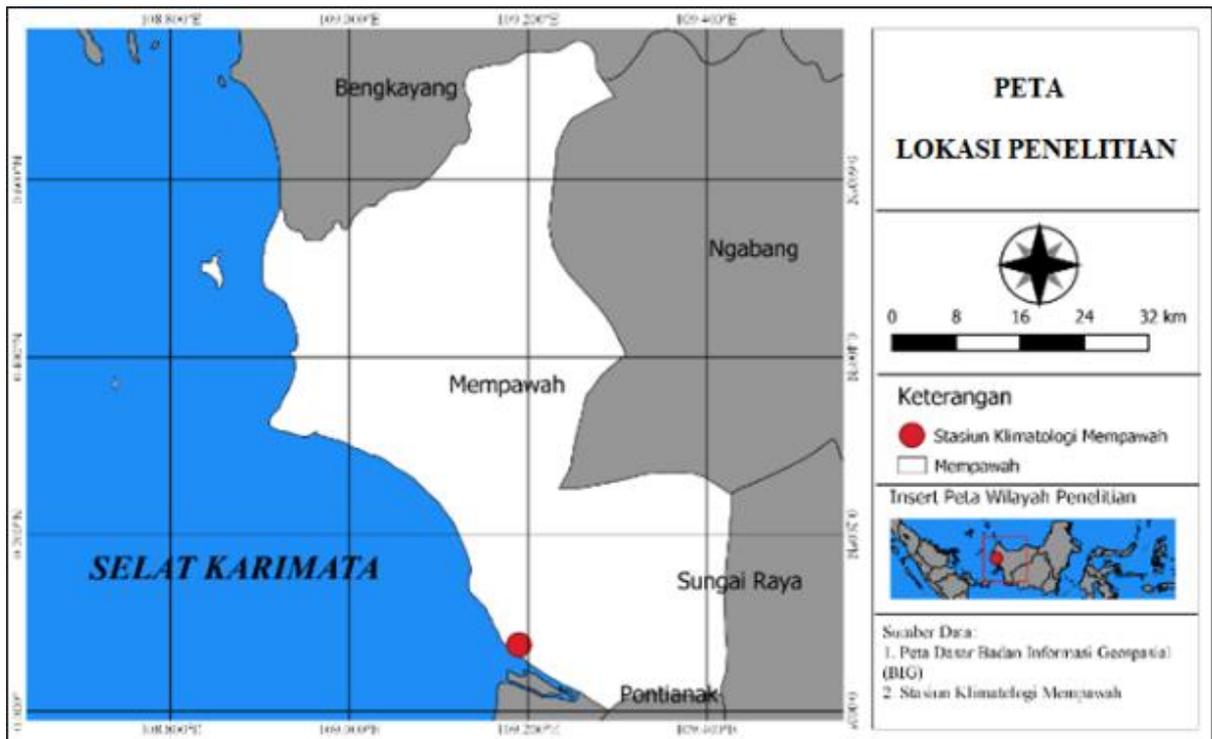


Gambar 1. Support Vector Machine

Sumber: Meyer dan Wien (Meyer & Wien, 2015)

Selain untuk klasifikasi, SVM juga dapat digunakan untuk regresi (Sánchez, Nieto, Fernández, del Coz Díaz, & Iglesias-Rodríguez, 2011). Variasi dari SVM yang dapat digunakan untuk regresi adalah *Support Vector Regression* (SVR) karena mampu memberikan model regresi untuk data non-linier (De Paz, Pérez, González, Corchado, & Corchado, 2010).

Guna mengetahui kualitas udara dan mengurangi dampak kesehatan akibat paparan polusi udara di Mempawah, Provinsi Kalimantan Barat maka pada penelitian ini digunakan metode SVR untuk menduga konsentrasi PM₁₀. Tujuan dari penelitian ini adalah melakukan estimasi PM₁₀ menggunakan SVR dan verifikasi dengan data observasi PM₁₀ di Mempawah periode tahun 2016 – 2019.



Gambar 2. Peta Lokasi Penelitian

2. METODE PENELITIAN

2.1. Lokasi Penelitian dan Data

Lokasi dalam penelitian ini adalah Kabupaten Mempawah yang merupakan salah satu kabupaten di Provinsi Kalimantan Barat. Lokasi dalam penelitian ditunjukkan pada Gambar 2.

Data yang digunakan pada penelitian ini didapat dari Stasiun Klimatologi Mempawah yang terletak di Kabupaten Mempawah. Data-data yang digunakan terdiri dari data rata-rata observasi harian konsentrasi PM₁₀, data pengamatan harian curah hujan, kelembaban, suhu, tekanan permukaan laut dan kecepatan angin sejak 1 Januari 2016 hingga 31 Desember 2019.

2.2. Normalisasi Data

Normalisasi data yang digunakan adalah *z-score* yang merupakan nilai standar dalam menentukan lokasi persis sebuah nilai dalam distribusi dengan menjelaskan jaraknya dari rata-rata dalam hal unit standar deviasi. Metode ini didasarkan pada nilai rata-rata dan standar deviasi dari data. Pada penelitian ini *z-score* akan digunakan untuk menormalisasi data kelembaban, suhu, tekanan permukaan laut dan kecepatan angin. Berikut rumus dari *z-score* ditunjukkan pada persamaan (1) (Chumney, 2019):

$$z - score = \frac{x - \bar{x}}{\sigma} \quad (1)$$

Keterangan:

x : nilai suatu data
 \bar{x} : nilai rata-rata
 σ : standar deviasi

2.3. Kernel Radial Basis Function

Fungsi *kernel* atau yang sering disebut dengan trik *kernel* dalam SVR digunakan untuk mentransformasi ruang *input* ke dalam fitur dimensi tinggi (Karatzoglou, Meyer, & Hornik, 2006). *Kernel Radial Basis Function* (RBF) merupakan salah satu *kernel* yang umum digunakan. *Kernel* ini memiliki parameter yang dapat memengaruhi kompleksitas pemilihan model yaitu parameter C dan gamma (Hsu, Chang, & Lin, 2003). Berikut fungsi *kernel* RBF ditunjukkan pada persamaan (2), (3) dan (4) (Ceperic, Ceperic, & Baric, 2013):

$$K(v, z) = \langle \Phi(x), \Phi(x') \rangle y \quad (2)$$

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (3)$$

$$= \exp(-\gamma (\langle x, x \rangle + \langle x', x' \rangle - 2\langle x, x' \rangle)), \text{ untuk } \gamma > 0 \quad (4)$$

2.4. Optimalisasi Parameter

Pada proses memperkirakan fungsi regresi, optimalisasi parameter sangat diperlukan. Terdapat tiga parameter yang harus ditentukan: nilai ϵ -insensitive (ϵ),

parameter *cost* (C) dan parameter *kernel* (γ) (Vapnik, 2013). Teknik yang dapat digunakan untuk optimalisasi parameter tersebut adalah *cross-validation*. Selain itu, dalam regresi SVM banyak peneliti juga menggunakan pendekatan *grid search* (Ceperic et al., 2013). Optimalisasi parameter untuk melakukan *grid search* dalam R menggunakan fungsi `tune()` pada paket `e1071` (Karatzoglou et al., 2006).

2.5. Support Vector Regression

Support Vector Regression (SVR) merupakan variasi dari SVM dalam memperoleh model regresi dengan cara mengubah dimensi data (De Paz et al., 2010). Ide utama dari SVR adalah menemukan fungsi $f(x)$ yang dapat memprediksi nilai pengamatan y dengan benar, dari nilai input baru x , berdasarkan data set pelatihan yang diberikan (Weizhen et al., 2014). Formulasi regresi SVM yang akan digunakan dalam penelitian ini adalah Vapnik's ϵ -tube Support Vector Regression (ϵ -SVR). Tujuan ϵ -SVR adalah melakukan regresi linier dalam ruang fitur dimensi tinggi yang dibuat oleh fungsi kernel menggunakan ϵ -insensitive loss dan mencoba mengurangi kompleksitas model dengan meminimalkan koefisien model (Ceperic et al., 2013). Pada perangkat lunak R, implementasi SVM diperkenalkan dalam paket `e1071`. Paket `e1071` menyediakan antarmuka ke `libsvm` melalui fungsi `svm()` yang dilengkapi pula dengan fungsi visualisasi dan penyetelan parameter (Karatzoglou et al., 2006). Bentuk umum fungsi regresi dinyatakan pada persamaan (5) sebagai berikut (Vapnik, 2013):

$$f(x) = w^T \Phi(x) + b \quad (5)$$

Dengan x merupakan vektor input, w^T merupakan parameter bobot, $\Phi(x)$ merupakan hasil pemetaan dari input vektor x dari ruang fitur dimensi rendah ke dimensi tinggi dan b merupakan bias.

2.6. Verifikasi Hasil Estimasi Konsentrasi PM₁₀

Verifikasi pada penelitian ini menggunakan korelasi Pearson, RMSE, bias dan *skill score* sebagai berikut:

1. Korelasi pearson

Korelasi adalah istilah statistik untuk mengukur kekuatan dan arah hubungan linier antara dua variabel atau lebih. Pada penelitian ini, korelasi yang digunakan adalah korelasi pearson. Korelasi person memiliki rentang nilai dari 0 hingga 1, sementara tanda + atau - hanya

menunjukkan arah hubungan. Berikut rumus korelasi Pearson ditunjukkan pada persamaan (6) (Usman & Akbar, 2000):

$$r_{XY} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}} \quad (6)$$

Keterangan:

r_{XY} : korelasi antara variabel X dan Y
n : jumlah data
X : variabel meteorologi (CH, Suhu, RH, Kecepatan Angin, QFF)
Y : PM₁₀

2. Root Mean Square Error (RMSE)

Salah satu metode untuk mengukur tingkat akurasi hasil prakiraan suatu model yang sering digunakan adalah RMSE. RMSE mengukur standar deviasi dari perbedaan antara nilai yang diperkirakan dan yang diamati. Semakin baik hasil prakiraan model apabila nilai RMSE semakin mendekati 0. Rumus RMSE yang digunakan ditunjukkan pada persamaan (7) (CAMS, 2020):

$$RMSE = \sqrt{\frac{1}{n} \sum_i (F_i - O_i)^2} \quad (7)$$

Keterangan:

F_i : nilai prakiraan
 O_i : nilai observasi
N : jumlah data

3. Bias

Bias adalah perbedaan antara rata-rata populasi dari pengukuran atau hasil pengujian dan nilai sebenarnya. Bias mengarah kepada prakiraan yang terlalu rendah atau terlalu tinggi. Nilai bias memiliki rentang nilai dari $-\infty$ hingga ∞ , dengan 0 sebagai nilai sempurna yang berarti prakiraan yang dihasilkan baik. Prakiraan dikatakan *overestimate* apabila nilai bias lebih dari 0, sedangkan *underestimate* apabila nilai bias kurang dari 0. Berikut rumus dari bias ditunjukkan pada persamaan (8) (Walther & Moore, 2005):

$$Bias = \frac{1}{N} \sum_{i=1}^N (F_i - O_i) \quad (8)$$

Keterangan:

F_i : nilai prakiraan
 O_i : nilai observasi
N : jumlah data

4. Skill score

Skill score adalah keakuratan prakiraan terhadap keakuratan prakiraan yang dihasilkan dari beberapa referensi seperti klimatologi. Nilai *skill* didasarkan pada *mean square error* (MSE) dan peka terhadap bias serta kesalahan dalam varian. *Skill score* memiliki rentang nilai dari $-\infty$ hingga 1. Apabila *skill score* bernilai positif maka keakuratan prakiraan lebih besar dari keakuratan prakiraan referensi, begitu juga sebaliknya ketika *skill score* bernilai negatif. Berikut rumus dari *skill score* ditunjukkan pada persamaan (9) (Murphy, 1988):

$$SS_{MSE} = \frac{MSE - MSE_{ref}}{MSE_{perf} - MSE_{ref}} = 1 \quad (9)$$
$$\frac{MSE}{MSE_{ref}}$$

Keterangan:

MSE : MSE prakiraan

MSE_{ref} : MSE acuan / klimatologi

3. HASIL DAN PEMBAHASAN

3.1. Estimasi Konsentrasi PM₁₀ di Mempawah Menggunakan SVR

Estimasi konsentrasi PM₁₀ di Mempawah periode tahun 2016 – 2019 menggunakan SVR berdasarkan enam parameter yaitu PM₁₀, curah hujan, kelembaban relatif (RH), suhu udara, tekanan permukaan laut dan kecepatan angin. Data yang digunakan dibagi menjadi dua periode yakni periode latihan (*training*) dari 1 Januari 2016 hingga 31 Desember 2018 dan periode uji dari 1 Januari hingga 31 Desember 2019.

Parameter yang digunakan dalam estimasi konsentrasi PM₁₀ dengan SVR pada periode latihan adalah $\epsilon=0,1$, $C=1$ dan $\gamma=0,2$. Hasil estimasi konsentrasi PM₁₀ yang ditunjukkan pada Gambar 3 dapat diketahui bahwa konsentrasi PM₁₀ memiliki variabilitas yang hampir sama dengan data observasinya, namun model belum dapat menduga PM₁₀ yang memiliki konsentrasi di atas 50 $\mu\text{g}/\text{m}^3$. Selanjutnya model estimasi di validasi dan didapatkan nilai korelasi positif atau arah korelasi searah sebesar 0,54. Sedangkan nilai RMSE sebesar 10,78 atau 0,86 kali standar deviasi observasi. Hal tersebut menunjukkan hasil konsentrasi PM₁₀ pada periode latihan memiliki nilai yang hampir mendekati konsentrasi PM₁₀ observasinya.

Agar mendapatkan model yang lebih baik, maka dilakukan optimalisasi parameter pada periode latihan. Rentang nilai yang diberikan

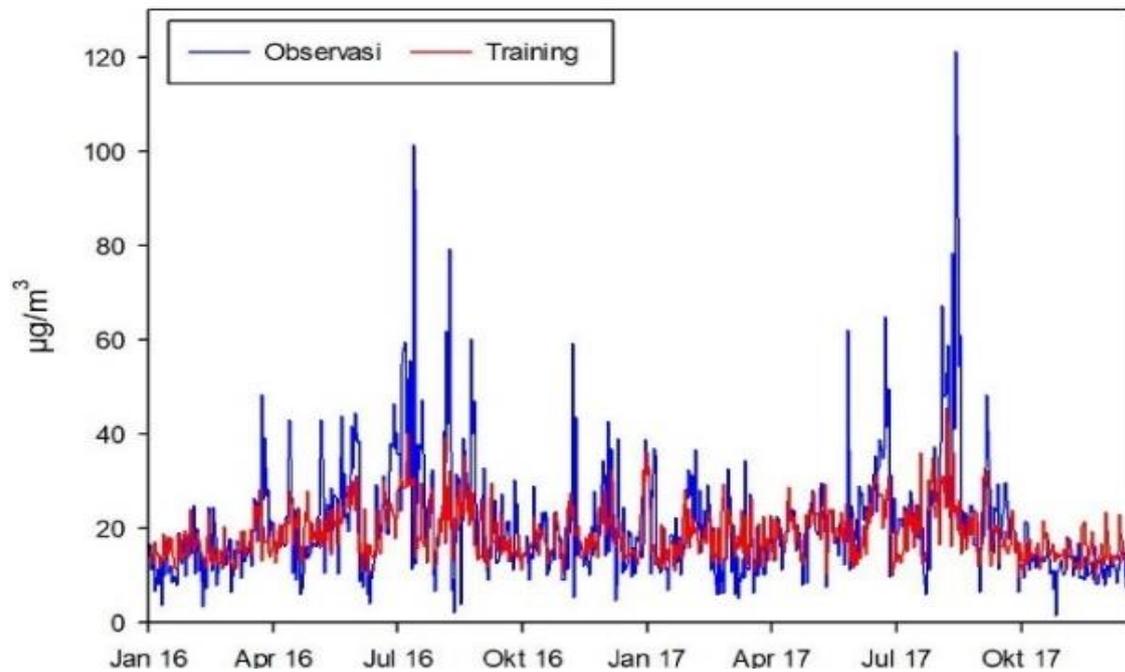
pada C dari 0 hingga 100 serta γ dari 0 hingga 1. Berdasarkan hasil evaluasi kinerja terhadap 1.100 model maka didapatkan nilai parameter optimal $C=1$ dan $\gamma=0,1$.

Parameter optimal yaitu $\epsilon=0,1$, $C=1$ dan $\gamma=0,1$ yang telah didapatkan digunakan pada model estimasi konsentrasi PM₁₀ periode latihan kembali. Hasil estimasi konsentrasi PM₁₀ menggunakan parameter optimal pada Gambar 4 tidak terlihat perbedaan yang signifikan dengan model estimasi sebelumnya. Selain itu, model juga masih belum dapat menduga konsentrasi PM₁₀ di atas 50 $\mu\text{g}/\text{m}^3$. Selanjutnya model estimasi kembali di validasi dan didapatkan nilai korelasi positif atau arah korelasi searah sebesar 0,51, sedangkan nilai RMSE sebesar 11 atau 0,88 kali standar deviasi observasi. Hal tersebut juga menunjukkan bahwa nilai estimasi konsentrasi PM₁₀ hampir mendekati nilai observasinya.

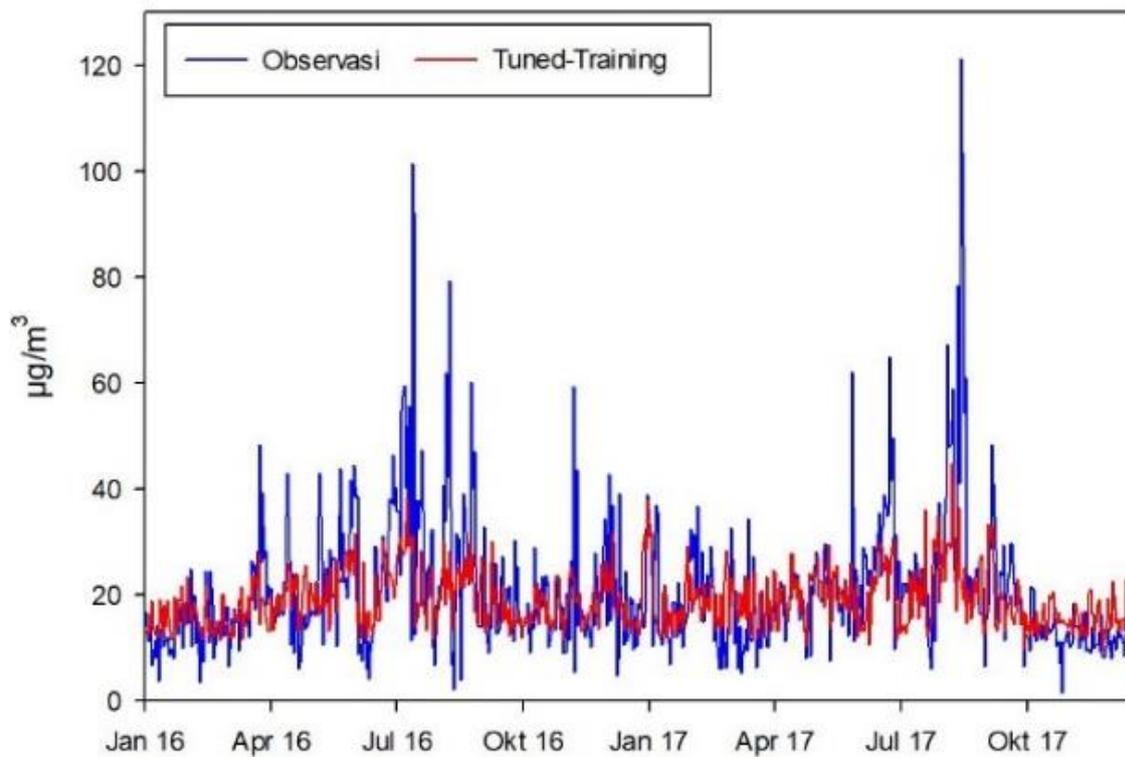
Berdasarkan hasil model estimasi konsentrasi PM₁₀ pada periode latihan baik sebelum maupun sesudah menggunakan parameter optimal dapat diketahui bahwa nilai konsentrasi PM₁₀ hampir mendekati nilai observasinya, tetapi SVR belum dapat menduga konsentrasi PM₁₀ yang memiliki konsentrasi tinggi di atas 50 $\mu\text{g}/\text{m}^3$. Hal tersebut dikarenakan adanya kejadian kebakaran hutan dan lahan di Mempawah. Hasil analisis spasial sebaran *hotspot* dari 1 Januari 2016 hingga 31 Desember 2019 menunjukkan bahwa Provinsi Kalimantan Barat merupakan salah satu wilayah dengan sebaran *hotspot* paling tinggi (Arisman, 2020).

Model yang telah dibangun selanjutnya dievaluasi menggunakan periode uji. Model estimasi periode uji menggunakan parameter $\epsilon=0,1$; $C=1$; dan $\gamma=0,1$. Hasil model estimasi ditampilkan dalam tiga periode yaitu Januari – Desember, Januari – Juli dan Agustus – Desember.

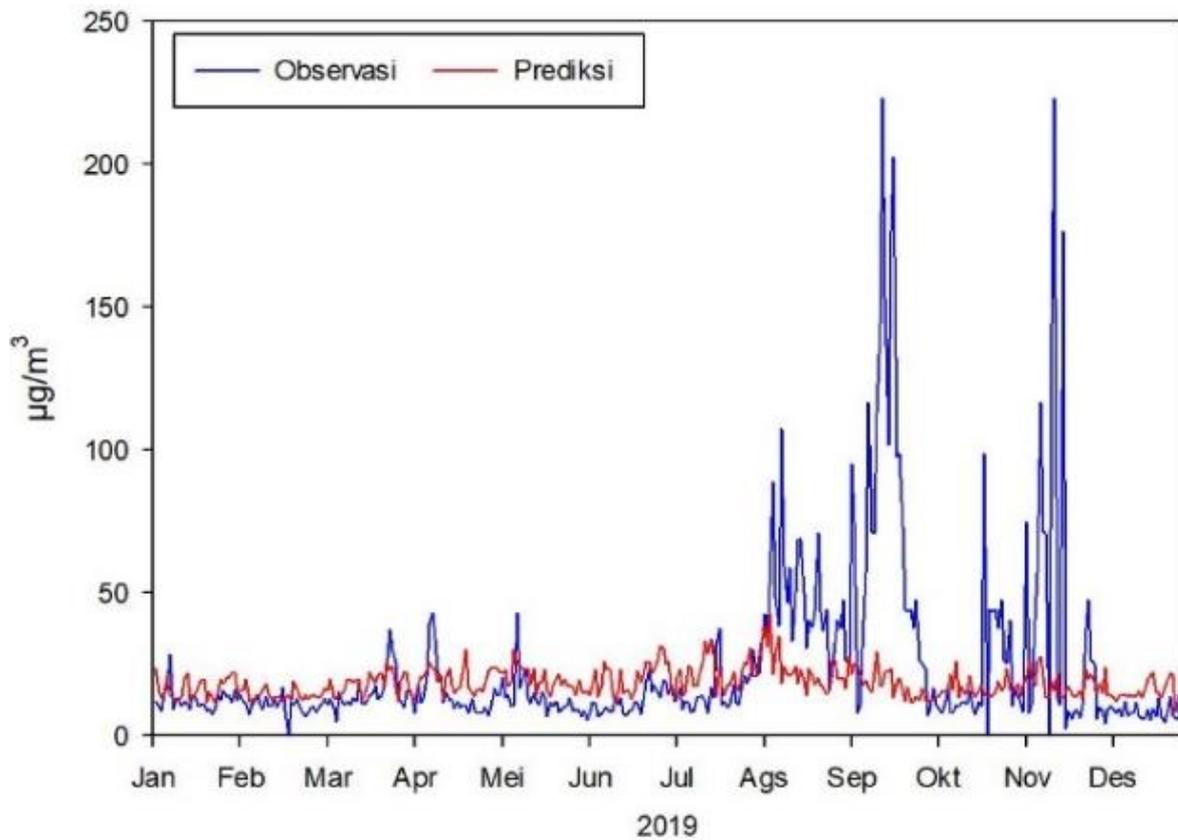
Pada periode 1 Januari hingga 31 Desember 2019 konsentrasi PM₁₀ menunjukkan rata-rata sebesar 18,56 $\mu\text{g}/\text{m}^3$, nilai minimum sebesar 8,27 $\mu\text{g}/\text{m}^3$ dan nilai maksimum sebesar 43,04 $\mu\text{g}/\text{m}^3$. Berdasarkan Gambar 5 dapat diketahui bahwa pada bulan-bulan awal tahun 2019 estimasi konsentrasi PM₁₀ memiliki variabilitas konsentrasi yang hampir sama dengan observasinya. Tetapi, pada bulan-bulan akhir tahun 2019 dimulai pada bulan Agustus terlihat bahwa estimasi konsentrasi PM₁₀ memiliki perbedaan yang signifikan dengan observasinya.



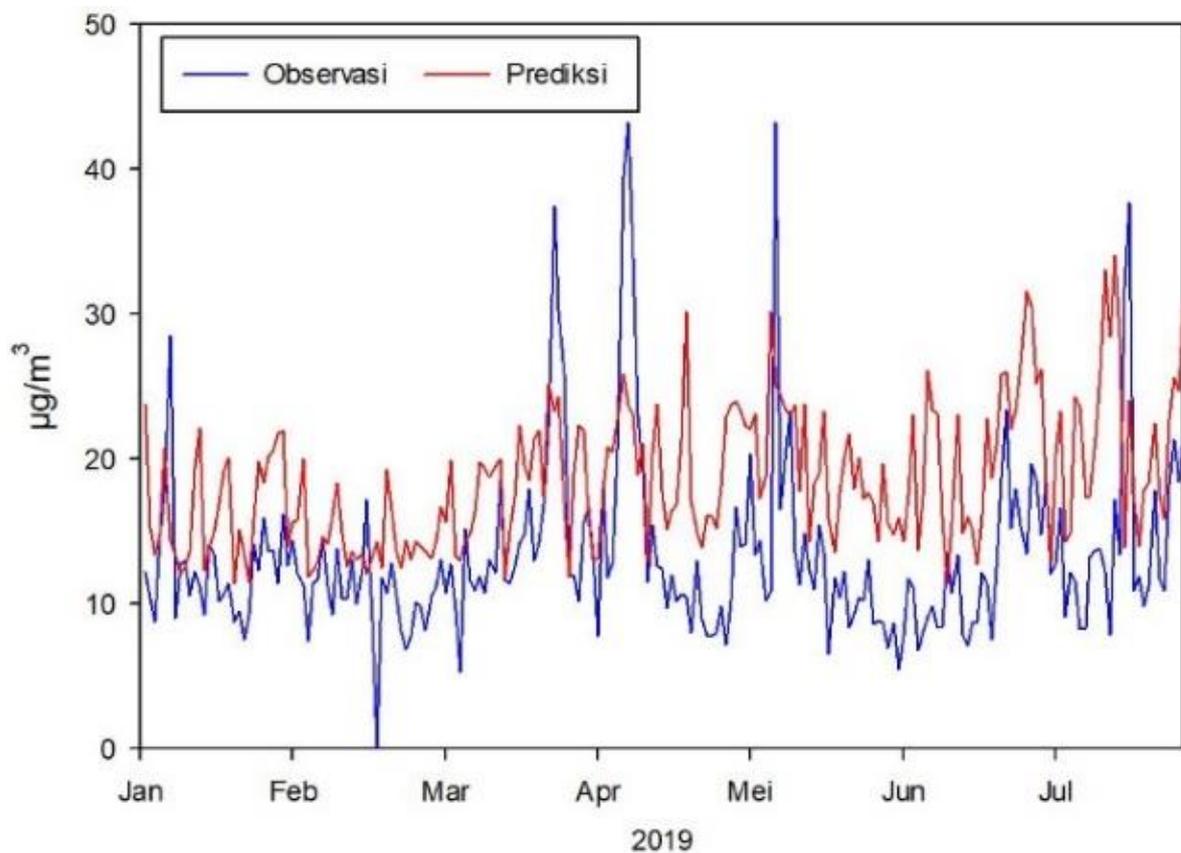
Gambar 3. Hasil Model Estimasi Konsentrasi PM₁₀ di Mempawah Periode Latih



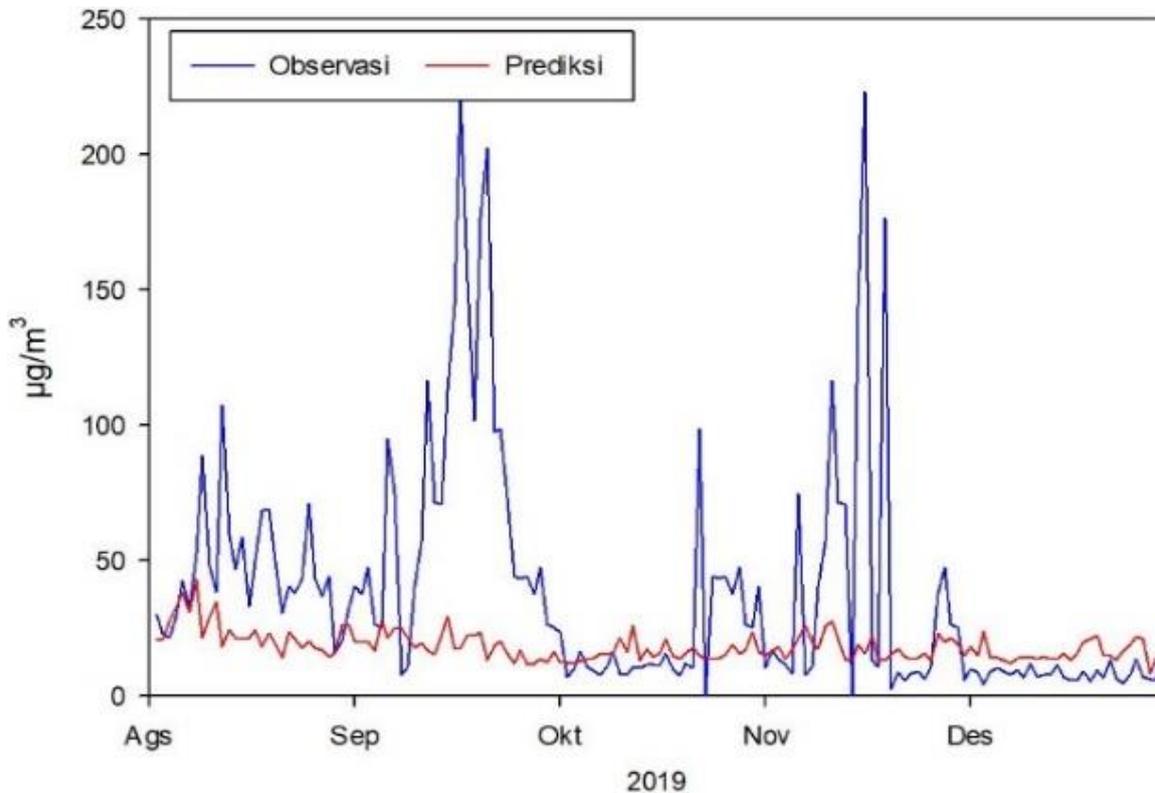
Gambar 4. Hasil Model Estimasi Konsentrasi PM₁₀ di Mempawah Periode Latih Dengan Parameter Optimal



Gambar 5. Hasil Model Estimasi Konsentrasi PM₁₀ di Mempawah Periode Januari – Desember 2019



Gambar 6. Hasil Model Estimasi Konsentrasi PM₁₀ di Mempawah Periode Januari – Juli 2019



Gambar 7. Hasil Model Estimasi Konsentrasi PM₁₀ di Mempawah Periode Agustus – Desember 2019

Periode 1 Januari hingga 31 Juli 2019 menunjukkan rata-rata konsentrasi PM₁₀ sebesar 18,64 µg/m³, nilai minimum sebesar 11,38 µg/m³ dan nilai maksimum sebesar 34,11 µg/m³. Terlihat pada Gambar 6 bahwa bulan Januari hingga Juli 2019 memiliki variabilitas konsentrasi PM₁₀ yang hampir sama dengan observasinya. Estimasi konsentrasi PM₁₀ menunjukkan hasil yang stabil dan setiap bulannya konsentrasi berada di bawah 50 µg/m³. Hal tersebut didukung pula dengan kejadian kebakaran hutan dan lahan selama periode 2016 hingga 2019 bahwa kejadian kebakaran hutan dan lahan di sebagian besar provinsi di Indonesia paling rendah terjadi di bulan Januari (Arisman, 2020).

Konsentrasi PM₁₀ periode Agustus hingga Desember 2019, menunjukkan rata-rata konsentrasi PM₁₀ sebesar 18,44 µg/m³, nilai minimum sebesar 8,27 µg/m³ dan nilai maksimum sebesar 43,04 µg/m³. Dapat dilihat bahwa pada Gambar 7, estimasi konsentrasi PM₁₀ periode Agustus hingga Desember 2019 memiliki hasil yang sangat jauh berbeda dengan observasinya. Perbedaan yang sangat signifikan terjadi pada bulan September dan November dikarenakan tingginya konsentrasi PM₁₀ observasi akibat dari kebakaran hutan dan lahan. Sementara pada bulan Oktober dan Desember 2019, hasil estimasi konsentrasi

PM₁₀ kembali memiliki fluktuasi yang hampir sama dengan observasinya. Meningkatnya konsentrasi PM₁₀ tersebut diakibatkan oleh kegiatan pembakaran hutan untuk pembukaan dan konversi lahan. Selain itu, di Kalimantan kegiatan pembakaran hutan dimulai pada awal musim kemarau dan mencapai puncak selama September dan Oktober 2019. Konsentrasi PM₁₀ akan mengalami penurunan ketika kegiatan pembakaran bertahap mereda dan berhenti pada bulan Oktober atau November 2019 ketika masuknya musim hujan (Heil & Goldammer, 2001).

Berdasarkan hal tersebut dapat diketahui bahwa, model estimasi konsentrasi PM₁₀ menggunakan SVR dengan parameter cuaca dapat memodelkan variabilitas konsentrasi PM₁₀ di Mempawah dengan cukup baik. Namun, terdapat beberapa kondisi tertentu saat SVR memerlukan tambahan parameter lingkungan yang mendukung kondisi tersebut sehingga dapat meningkatkan kualitas dari hasil estimasi konsentrasi PM₁₀ menjadi lebih baik seperti di saat terjadi kebakaran hutan dan lahan. Hal tersebut diperkuat pula dengan pengaruh faktor iklim selalu mendukung terjadinya kebakaran hutan dan lahan di Kalimantan, tapi pengaruh faktor antropogenik lebih dominan daripada faktor alami ketika kejadian dengan luas kebakaran

hutan dan lahan yang tinggi (Mareta, Hidayat, Hidayati, & Latifah, 2020).

3.1. Verifikasi Hasil Estimasi Konsentrasi PM₁₀ Menggunakan SVR Dengan Data Observasi di Mempawah

Tabel 1. Verifikasi Model Estimasi Konsentrasi PM₁₀ di Mempawah di Tahun 2019

Verifikasi i	Januari – Desembe r	Januar i – Juli	Agustus – Desembe r
Korelasi	0,14	0,39	0,17
RMSE	31,41	8,13	47,56
Bias	-5,42	5,31	-20,30
<i>skill</i> <i>score</i>	-0,01	0,56	-0,07

Nilai verifikasi model estimasi konsentrasi PM₁₀ di Mempawah tahun 2019 ditunjukkan pada Tabel 1. Korelasi ketiga periode hasil model estimasi konsentrasi PM₁₀ dengan observasinya menunjukkan korelasi positif atau searah walaupun dengan nilai yang cukup kecil. Nilai korelasi tertinggi terdapat pada periode Januari hingga Juli 2019. Untuk RMSE, periode Januari hingga Desember 2019 memiliki nilai yang hampir sama dengan nilai standar deviasi observasinya yaitu 32,34. Periode Januari hingga Juli 2019 memiliki nilai RMSE yang lebih besar dari standar deviasi observasi atau 1,32 kali standar deviasi observasinya. Sementara itu, periode Agustus hingga Desember 2019 memiliki nilai RMSE yang juga lebih besar dari standar deviasi observasi atau 1,09 kali standar deviasi observasinya. Berdasarkan hal tersebut dapat diketahui bahwa nilai RMSE ketiga periode memiliki nilai yang tidak jauh berbeda dengan observasinya, Hal tersebut menunjukkan bahwa konsentrasi PM₁₀ hasil model estimasi memiliki nilai yang hampir mirip dengan nilai konsentrasi PM₁₀ observasi.

Bias yang dihasilkan model estimasi konsentrasi PM₁₀ periode Januari hingga Desember 2019 dan Agustus hingga Desember 2019 bernilai negatif. Hal tersebut berarti bahwa estimasi rata-rata konsentrasi PM₁₀ lebih rendah daripada rata-rata konsentrasi PM₁₀ observasinya (*underestimate*). Sedangkan bias pada periode Januari hingga Juli 2019 bernilai positif yang berarti estimasi rata-rata konsentrasi PM₁₀ lebih tinggi daripada rata-rata konsentrasi PM₁₀ observasinya (*overestimate*). Selain itu, verifikasi model menggunakan *skill score* pada periode Januari hingga Desember 2019 dan Agustus hingga Desember 2019

menunjukkan akurasi model dalam menduga konsentrasi PM₁₀ berdasarkan klimatologinya masih belum baik. Sementara, *skill score* periode Januari hingga Juli 2019 menunjukkan akurasi model yang baik dalam menduga konsentrasi PM₁₀ berdasarkan klimatologinya.

4. KESIMPULAN

Estimasi konsentrasi PM₁₀ di Mempawah menggunakan SVR pada periode latih baik sebelum maupun sesudah menggunakan parameter optimal menunjukkan variabilitas yang hampir sama dengan konsentrasi PM₁₀ observasi.

Estimasi konsentrasi PM₁₀ di Mempawah menggunakan SVR pada periode uji dapat memodelkan konsentrasi PM₁₀ dengan baik, hanya saja tidak di saat terjadi kondisi-kondisi tertentu. Hal tersebut terlihat pada periode Januari hingga Juli 2019 yang memiliki fluktuasi konsentrasi PM₁₀ yang hampir sama dengan observasinya dan menunjukkan nilai konsentrasi yang tidak jauh dari rata-rata setiap bulannya.

Verifikasi luaran model estimasi konsentrasi PM₁₀ dengan nilai observasinya menunjukkan bahwa model estimasi konsentrasi PM₁₀ di Mempawah menggunakan SVR dengan parameter cuaca dapat memberikan hasil estimasi konsentrasi PM₁₀ dengan baik. Hal tersebut dapat dilihat dari nilai RMSE yang dihasilkan model estimasi hampir sama dengan nilai standar deviasi observasi. Selain itu, akurasi model estimasi pada periode Januari hingga Juli 2019 sudah cukup baik.

UCAPAN TERIMA KASIH

Ucapan terima kasih kepada Badan Meteorologi, Klimatologi dan Geofisika khususnya Stasiun Klimatologi Mempawah, Kalimantan Barat atas dukungan data dalam penelitian ini.

DAFTAR PUSTAKA

- Arisman, A. (2020). Analisis Tren Kebakaran Hutan dan Lahan di Indonesia Periode Tahun 2015-2019. *JURNAL SAINS TEKNOLOGI & LINGKUNGAN*, 6(1), 1-9.
- CAMS, C. A. M. S. (2020). Verification plots: documentation. Retrieved from <http://macc-raq-op.meteo.fr/doc/userguideverificationstatistics.pdf>
- Ceperic, E., Ceperic, V., & Baric, A. (2013). A strategy for short-term load

- forecasting by support vector regression machines. *IEEE Transactions on Power Systems*, 28(4), 4356-4364.
- Chantara, S. (2012). PM10 and its chemical composition: a case study in Chiang Mai, Thailand. *Air quality-monitoring and modeling*, 205-230.
- Chumney, F. (2019). z-Scores. Retrieved from https://www.westga.edu/academics/research/vrc/assets/docs/zScores_HA_NDOUT.pdf
- De Paz, J. F., Pérez, B., González, A., Corchado, E., & Corchado, J. M. (2010). A support vector regression approach to predict carbon dioxide exchange. In *Distributed Computing and Artificial Intelligence* (pp. 157-164): Springer.
- EPA, E. P. A. (2019). Particulate Matter (PM) in New England. Retrieved from <https://www3.epa.gov/region1/airquality/partic.html>
- Heil, A., & Goldammer, J. (2001). Smoke-haze pollution: a review of the 1997 episode in Southeast Asia. *Regional Environmental Change*, 2(1), 24-37.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification" (<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>).
- Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support vector machines in R. *Journal of statistical software*, 15(9), 1-28.
- Lu, W.-Z., & Wang, W.-J. (2005). Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. *Chemosphere*, 59(5), 693-701.
- Lu, W., Wang, W., Fan, H., Leung, A., Xu, Z., Lo, S., & Wong, J. (2002). Prediction of pollutant levels in causeway bay area of Hong Kong using an improved neural network model. *Journal of environmental engineering*, 128(12), 1146-1157.
- Mareta, L., Hidayat, R., Hidayati, R., & Latifah, A. L. (2020). Pengaruh Faktor Alami dan Antropogenik Terhadap Luas Kebakaran Hutan dan Lahan di Kalimantan. *Jurnal Tanah dan Iklim*, 43(2), 147-155.
- Martin, E. P. (2011). *Comparative Performance of Different Statistical Models for Predicting Ground-Level Ozone (O3) and Fine Particulate Matter (PM2.5) Concentrations in Montréal, Canada*. Concordia University,
- Meilantika, A. D., Hadisaputro, S., & Setiawati, M. (2018). *Faktor Risiko Host dan Environment yang Berpengaruh Terhadap Kejadian Pneumonia Pada Balita (Studi di Wilayah Kerja Puskesmas Rawat Jalan Wajok Hulu Kecamatan Siantan Kabupaten Mempawah)*. School of Postgraduate,
- Meyer, D., & Wien, F. T. (2015). Support vector machines. *The Interface to libsvm in package e1071*, 28.
- Munir, M., Akbar, A. R., Badaruddin, B., & Wahdah, R. (2018). Hubungan Cuaca dan Konsentrasi Pm10 (Studi Kasus di Kota Banjarbaru). *EnviroScienteeae*, 14(1), 46-61.
- Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly weather review*, 116, 2417-2424.
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). Support vector machine teori dan aplikasinya dalam bioinformatika. *Kuliah Umum Ilmu Komputer. com*.
- RRI.co.id. (2019). Karhutla Mempawah Sudah Memakan Lahan Seluas 473 Hektar. Retrieved from http://rri.co.id/post/berita/705918/mitigasi_bencana/karhutla_mempawah_sudah_memakan_lahan_seluas_473_hektar.html
- Sánchez, A. S., Nieto, P. G., Fernández, P. R., del Coz Díaz, J., & Iglesias-Rodríguez, F. J. (2011). Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). *Mathematical and Computer Modelling*, 54(5-6), 1453-1466.
- Suarapemredkalbar.com. (2019). Kualitas Udara Mempawah Berbahaya. Retrieved from <https://www.suarapemredkalbar.com/berita/mempawah/2019/09/19/kualitas-udara-mempawah-berbahaya>

- Tacconi, L. (2003). *Kebakaran hutan di Indonesia: penyebab, biaya dan implikasi kebijakan*: CIFOR.
- Ul-Saufie, A. Z., Yahaya, A. S., Ramli, N. A., Rosaida, N., & Hamid, H. A. (2013). Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmospheric Environment*, 77, 621-630.
- Usman, H., & Akbar, R. P. S. (2000). *Pengantar Statistika*. Jakarta: Bumi Aksara.
- Vapnik, V. (2013). *The nature of statistical learning theory*: Springer science & business media.
- Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28, 815-829.
- Wang, W., Men, C., & Lu, W. (2008). Online prediction model based on support vector machine. *Neurocomputing*, 71(4-6), 550-558.
- Wang, X., & Zhong, Y. (2003). *Statistical learning theory and state of the art in SVM*. Paper presented at the The Second IEEE International Conference on Cognitive Informatics, 2003. Proceedings.
- Weizhen, H., Zhengqiang, L., Yuhuan, Z., Hua, X., Ying, Z., Kaitao, L., . . . Yan, M. (2014). *Using support vector regression to predict PM10 and PM2. 5*. Paper presented at the IOP Conference Series: Earth and Environmental Science.
- WHO, W. H. O. (2003). Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide: Report on a WHO Working Group. Retrieved from <https://apps.who.int/iris/handle/10665/107478>
- Yulianti, N. (2018). *Pengenalan bencana kebakaran dan kabut asap lintas batas (studi kasus eks proyek lahan gambut sejuta hektar)*: Bogor: IPB Press.