



# Prediksi Spasial Kerapatan Vegetasi Perkotaan dengan Pendekatan Algoritma Time Series Untuk Mendukung Pertumbuhan Ekonomi Hijau

Yudistira Bagus Pratama<sup>1,\*</sup>, Nurzaidah Putri Dalimunthe<sup>2</sup>, Mega Sukma<sup>3</sup>

<sup>1</sup>Fakultas Teknik dan Sains, Program Studi Ilmu Komputer, Universitas Muhammdiyah Bangka Belitung, Pangkal Pinang  
Jl. KH A Dahlan No.Km.4, Keramat, Kec. Rangkui, Kota Pangkal Pinang, Kepulauan Bangka Belitung

<sup>2</sup>Fakultas Teknik dan Sains, Program Studi Konservasi Sumberdaya Alam, Universitas Muhammdiyah Bangka Belitung,  
Pangkal Pinang

Jl. KH A Dahlan No.Km.4, Keramat, Kec. Rangkui, Kota Pangkal Pinang, Kepulauan Bangka Belitung

<sup>3</sup>Fakultas Teknik dan Sains, Program Studi Teknik Sipil, Universitas Muhammdiyah Bangka Belitung, Pangkal Pinang  
Jl. KH A Dahlan No.Km.4, Keramat, Kec. Rangkui, Kota Pangkal Pinang, Kepulauan Bangka Belitung

Email: <sup>1\*</sup>yudistira.bagus@unmuhbabel.ac.id, <sup>2</sup>nurzaidah.putridalimunthe@unmuhbabel.ac.id,

<sup>3</sup>mega.sukma@unmuhbabel.ac.id

Email Penulis Korespondensi: yudistira.bagus@unmuhbabel.ac.id

Submitted: 12/11/2024; Accepted: 31/01/2025; Published: 31/01/2025

**Abstrak**–Urgensi penelitian ini berdasarkan data Dinas Kependudukan dan Pencatatan Sipil Kota Pangkalpinang pada tahun 2020, jumlah penduduknya mencapai 218.569 jiwa lalu terus bertambah menjadi 232.915 jiwa pada tahun 2023. Pentingnya pemantauan kerapatan vegetasi dalam konteks pertumbuhan ekonomi hijau, yang menuntut evaluasi yang cermat terhadap keseimbangan antara pembangunan ekonomi dan pelestarian lingkungan. Dengan pertumbuhan urban yang cepat, pemerintah kota Pangkalpinang membutuhkan pendekatan yang mutakhir untuk memprediksi kerapatan vegetasi dengan akurat. Dengan mengidentifikasi faktor-faktor yang mempengaruhi kerapatan vegetasi lahan, penelitian ini bertujuan untuk mengembangkan model machine learning yang dapat memprediksi kondisi kerapatan vegetasi lahan dalam kurun waktu tertentu di masa depan. Metode penelitian ini melibatkan pengumpulan data spasial kerapatan vegetasi dalam rentang 11 tahun menggunakan teknologi remote sensing atau pemantauan jarak jauh, seperti citra satelit. Selanjutnya, data time series akan dianalisis dan dimodelkan menggunakan teknik machine learning, dengan fokus pada algoritma yang dapat mengatasi dinamika spasial dan temporal dari kerapatan vegetasi. Algoritma machine learning, terutama algoritma time series seperti Autoregresif Integrated Moving Average (ARIMA) akan digunakan untuk membangun model prediksi spasial kerapatan vegetasi. Hasil penelitian ini menunjukkan bahwa penggunaan ARIMA mampu menghasilkan model prediksi yang akurat dalam memproyeksikan kerapatan vegetasi di Kota Pangkalpinang. Model ARIMA menunjukkan kinerja yang kuat dengan metrik kesalahan yang rendah, yang menunjukkan efektivitasnya dalam membuat prediksi yang akurat untuk kumpulan data yang diberikan. Hasil dari penelitian ini diharapkan dapat memberikan informasi yang berharga bagi pemerintah kota Pangkalpinang dalam mengambil keputusan terkait pengelolaan lingkungan dan pembangunan ekonomi hijau. Dengan melibatkan kerjasama antara peneliti dengan tiga ilmu kepakaran yang saling melengkapi diantaranya ilmu komputer, teknik sipil dan konservasi sumberdaya alam serta pemangku kebijakan.

**Kata Kunci:** Pembelajaran Mesin; Pengindraan Jauh; Geospasial; NDVI; Ekonomi Hijau

**Abstract**–The urgency of this research is based on data from the Pangkalpinang City Population and Civil Registration Service in 2020, the population reached 218,569 people and continued to grow to 232,915 people in 2023. The importance of monitoring vegetation density in the context of green economic growth, which requires careful evaluation of the balance between economic development and environmental conservation. With rapid urban growth, the Pangkalpinang city government requires a variety of approaches to accurately predict vegetation density. By identifying the factors that affect land vegetation density, this study aims to develop a machine learning model that can predict land vegetation density conditions over a certain period of time in the future. This research method involves collecting spatial vegetation density data over a period of 11 years using remote sensing technology or remote monitoring, such as satellite imagery. Furthermore, time series data will be analyzed and modeled using machine learning techniques, focusing on algorithms that can overcome the spatial and temporal dynamics of vegetation density. Machine learning algorithms, especially time series algorithms such as Autoregressive Integrated Moving Average (ARIMA) will be used to build a spatial prediction model for vegetation density. The results of this study indicate that the use of ARIMA is able to produce an accurate prediction model in projecting vegetation density in Pangkalpinang City. The ARIMA model shows strong performance with low error metrics, indicating its effectiveness in making accurate predictions for the given data set. The results of this study are expected to provide valuable information for the Pangkalpinang city government in making decisions related to environmental management and green economic development. By involving collaboration between researchers with three complementary expertise including computer science, civil engineering and natural resource conservation and policy makers.

**Keywords:** Machine Learning; Remote Sensing; Geospatial; NDVI; Green Economy

## 1. PENDAHULUAN

Kota Pangkalpinang, sebagai pusat ekonomi di Kepulauan Bangka Belitung, menghadapi tantangan dalam menjaga keseimbangan alam antara pertumbuhan ekonomi yang pesat dan pelestarian lingkungan karena urbanisasi yang terus meningkat. Berdasarkan data Dinas Kependudukan dan Pencatatan Sipil Kota Pangkalpinang pada tahun 2020, jumlah penduduknya mencapai 218.569 jiwa lalu terus bertambah menjadi 232.915 jiwa pada tahun 2023. Semakin terdesaknya alokasi ruang untuk vegetasi di perkotaan menyebabkan kualitas lingkungan

menurun [1]. Perubahan penggunaan lahan yang tidak sesuai dengan kaidah-kaidah rencana tata ruang dapat mengakibatkan menurunnya kualitas lingkungan, degradasi kerusakan lingkungan serta berkurangnya sumber daya alam [2].

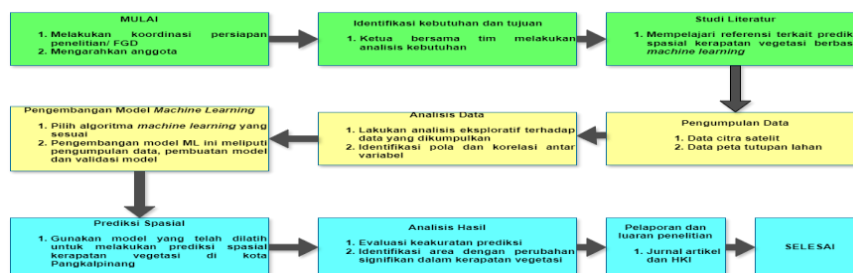
Vegetasi lahan memainkan peran penting dalam menjaga kualitas lingkungan dan mendukung pertumbuhan ekonomi hijau [3]. Vegetasi membantu membersihkan udara, mengurangi polusi suara, dan menyediakan ruang publik yang nyaman bagi masyarakat [4]. Tersedianya teknologi seperti drone, foto satelit, dan sensor jarak jauh menghasilkan data dalam jumlah besar, machine learning dapat membantu proses pengolahan data berdasarkan sumber data tersebut menjadi lebih efektif [5]. Salah satu cara yang efektif untuk mengetahui perubahan dalam kepadatan vegetasi di kota Pangkalpinang adalah dengan melakukan pengamatan dengan teknologi remote sensing atau penginderaan jauh yang dapat dipantau melalui Citra Landsat 8 Collection 2 level 2 dengan surface reflectance bands 30 meter [6]. Indeks vegetasi yang digunakan untuk mengetahui kepadatan vegetasi yaitu nilai Normalized Difference Vegetation Index (NDVI) [7]. NDVI merupakan suatu transformasi untuk menonjolkan aspek vegetasi sehingga dapat menunjukkan tingkat kepadatan vegetasi yang ada di lapangan [8].

Studi ilmiah tentang algoritma dan model statistik yang digunakan sistem komputer untuk melakukan tugas tertentu seperti prediksi dikenal sebagai machine learning [9]. Penurunan kepadatan vegetasi dapat membawa dampak negatif yang signifikan terhadap lingkungan dan kualitas hidup masyarakat [10]. Dengan mengidentifikasi faktor-faktor yang mempengaruhi kepadatan vegetasi lahan berbasis machine learning, penelitian ini bertujuan untuk mengembangkan model machine learning yang dapat memprediksi kondisi kepadatan vegetasi lahan dalam kurun waktu tertentu di masa depan sehingga dapat memberikan kontribusi dan insight untuk pemerintah kota Pangkalpinang dalam merumuskan kebijakan yang tepat dalam mendukung pertumbuhan ekonomi hijau di kota tersebut. Sebagai upaya untuk meningkatkan kepadatan vegetasi dan mendukung pertumbuhan ekonomi hijau maka penelitian ini melibatkan peneliti dengan tiga ilmu kepakaran yang saling melengkapi yaitu ilmu komputer, konservasi sumberdaya alam dan teknik sipil serta kerjasama dengan pemangku kebijakan yaitu instansi terkait di pemerintah kota Pangkalpinang.

State of art (SoA) merupakan penelitian terdahulu yang digunakan sebagai panduan ataupun acuan dan perbandingan dalam melakukan penelitian. SoA penelitian ini antara lain penelitian [11] menggunakan metode klasifikasi terbimbing pada data citra penginderaan jauh, penelitian [12] fokus pada klasifikasi tutupan lahan pada citra sentinel-2 menggunakan algoritme random forest dan penelitian [13] yang melakukan klasifikasi wilayah potensi risiko kerusakan lahan akibat bencana tsunami menggunakan algoritma ANN. Berdasarkan State of art (SoA) sebelumnya dapat diambil kesimpulan penelitian saat ini juga berfokus pada pengembangan model machine learning untuk mendukung pertumbuhan ekonomi hijau. Algoritma prediksi spasial kepadatan vegetasi menggunakan time series menjadi bagian krusial dalam penelitian ini, karena pada penelitian sebelumnya hanya terbatas klasifikasi nilai kepadatan vegetasi terkini yang tidak dapat memprediksi nilai perubahan kepadatan vegetasi di masa depan, sehingga penelitian ini menghadirkan kebaruan dengan mengembangkan model prediksi kepadatan vegetasi berbasis algoritma time series, yaitu ARIMA, untuk memproyeksikan perubahan kepadatan vegetasi di masa depan secara akurat, mendukung pengambilan kebijakan yang berwawasan lingkungan. Penelitian ini selanjutnya dapat dikembangkan untuk optimalisasi penggunaan jasa lingkungan ekologi dengan modifikasi machine learning yang mendukung ekonomi hijau.

## 2. METODOLOGI PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah dengan memanfaatkan remote sensing [14] dan algoritma time series [15] [16]. Kedua teknologi ini dapat membantu dalam pengembangan model prediksi spasial kepadatan vegetasi untuk mendukung pertumbuhan ekonomi hijau yang berkelanjutan dan berwawasan lingkungan di kota Pangkalpinang. Teknologi remote sensing dapat membantu peneliti mengumpulkan informasi tentang objek atau daerah tanpa harus berada di lokasi fisik. Teknologi ini dapat membantu pemantauan lingkungan yang luas dan terus-menerus serta memungkinkan deteksi perubahan lingkungan seperti deforestasi, perubahan penggunaan lahan, erosi tanah, dan perubahan iklim [17].



Gambar 1. Diagram Alir Penelitian

Sedangkan, algoritma time series dapat membantu untuk menganalisis data historis untuk mengidentifikasi pola dan tren, dan kemudian menggunakan pola tersebut untuk memprediksi nilai masa depan



dalam deret waktu [18]. Hasil validasi analisis data dengan algoritma time series dapat digunakan dalam membuat rekomendasi kebijakan dan mengembangkan strategi yang tepat untuk meningkatkan kerapatan vegetasi di kawasan pemukiman kota Pangkalpinang untuk mencapai pembangunan ekonomi hijau yang berkelanjutan.

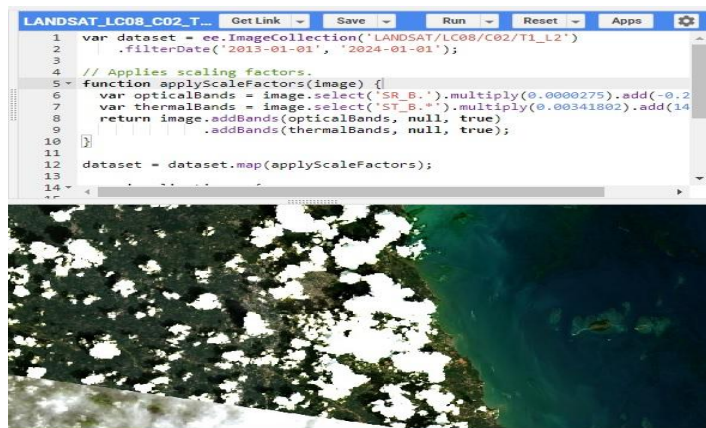
Berikut adalah tahapan metode pengembangan algoritma prediksi spasial berbasis machine learning yang dapat dilakukan untuk mendukung pertumbuhan ekonomi hijau dan pembangunan berkelanjutan yang berwawasan lingkungan di kota Pangkalpinang: 1). **Memulai Persiapan untuk penelitian:** Tahap ini meliputi pemilihan tim peneliti, pengumpulan data dan informasi, serta merumuskan tujuan dan masalah penelitian dengan jelas; 2). **Identifikasi kebutuhan dan tujuan:** Langkah mengidentifikasi kebutuhan dan tujuan dari prediksi spasial kerapatan vegetasi berbasis machine learning. Hal ini dilakukan untuk menentukan fokus dan arah pengembangan model machine learning yang akan dibuat. Selain itu melakukan analisis kebutuhan stakeholder, seperti peta perubahan lahan, perubahan iklim, dan pemerintah kota, untuk memahami masalah dan kebutuhan yang perlu diselesaikan oleh model machine learning prediksi spasial kerapatan vegetasi; 3). **Studi literatur:** Melakukan studi literatur untuk mengumpulkan informasi tentang teknologi remote sensing berbasis machine learning yang relevan dan aplikasinya dalam green industry; 4). **Pengumpulan data:** Mengumpulkan data spasial tentang kerapatan vegetasi Kota Pangkalpinang seperti citra satelit, peta tutupan lahan dan data terkait faktor yang mempengaruhi kerapatan vegetasi seperti perubahan iklim yang dapat membantu dalam pengembangan model machine learning prediksi spasial kerapatan vegetasi. Pendekatan validasi dengan skala faktor memastikan bahwa hasil analisis remote sensing konsisten dan relevan dengan aplikasi spesifik. Penyesuaian resolusi, normalisasi temporal, dan resampling sangat penting untuk mengurangi bias dan kesalahan yang disebabkan oleh perbedaan skala antara data remote sensing dan referensi.; 5). **Analisis Data:** Tahapan selanjutnya melakukan analisis eksploratif terhadap data yang berkaitan dengan kerapatan vegetasi di Kota Pangkalpinang. Data dapat diperoleh dari berbagai sumber, seperti data pemerintah, data industri, dan data satelit. Mengidentifikasi pola dan korelasi antar variabel data. Mengelola data yang hilang dengan linear interpolation menggunakan nilai sebelum dan setelah data hilang untuk memperkirakan nilai yang hilang. ; 6). **Pengembangan model machine learning:** Selanjutnya mengembangkan model machine learning yang sesuai dengan kebutuhan dan tujuan penelitian. Pengembangan model machine learning ini meliputi pengumpulan data, pembuatan model, dan validasi model dengan menggunakan algoritma time series. Model yang dikembangkan dapat mengintegrasikan berbagai sumber data, seperti data tutupan lahan, data perubahan iklim, dan data index vegetasi. Algoritma ARIMA (Autoregressive Integrated Moving Average) merupakan metode umum yang digunakan untuk memprediksi nilai masa depan dalam deret waktu (time series). Implementasi algoritma ARIMA dapat dibagi menjadi beberapa tahapan yaitu : 1. Kumpulkan data deret waktu yang ingin diprediksi. Pastikan data tersebut lengkap dan bebas dari missing values. 2. Lakukan analisis data awal untuk memahami karakteristik data, seperti tren, musimanitas, dan stasioneritas. 3. Gunakan metode identifikasi model, seperti Augmented Dickey-Fuller (ADF) test atau KPSS test, untuk menentukan tingkat differencing (d) yang diperlukan. 4. Tentukan nilai parameter autoregressive (p) dan moving average (q) dengan melihat autokorelasi (ACF) dan partial autocorrelation (PACF) plot. 5. Gunakan metode maximum likelihood estimation (MLE) untuk memperkirakan parameter model ARIMA yang dipilih pada tahap sebelumnya. 6. Hitung nilai akurasi model dengan menggunakan metrik statistik seperti Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), atau Mean Absolute Percentage Error (MAPE). 7. Lakukan analisis diagnostik untuk melihat apakah model ARIMA yang dipilih sudah sesuai dan tidak terdapat misspesifikasi. 8. Gunakan model ARIMA yang telah diestimasi untuk memprediksi nilai deret waktu di masa depan. 9. Hitung interval prediksi untuk menunjukkan tingkat ketidakpastian dalam prediksi. ; 7). **Prediksi spasial:** Tahapan selanjutnya melakukan prediksi spasial kerapatan vegetasi di Kota Pangkalpinang menggunakan model machine learning yang telah dilatih untuk mendapatkan prediksi terhadap perubahan nilai kerapatan vegetasi di masa depan; 8). **Analisis hasil:** Setelah prediksi spasial dilakukan, selanjutnya menganalisis hasil berupa evaluasi keakuratan prediksi apakah sudah mendapatkan hasil yang diharapkan lalu mengidentifikasi area dengan perubahan signifikan dalam kerapatan vegetasi lahan di Kota Pangkalpinang; 9). **Pelaporan dan luaran penelitian:** Langkah terakhir adalah pelaporan dan luaran penelitian. Melakukan pelaporan hasil penelitian dan menghasilkan luaran penelitian berupa artikel ilmiah dan luaran berupa HKI kode sumber program prediksi spasial kerapatan vegetasi untuk membagikan hasil penelitian dan memberikan masukan bagi pengembangan prediksi spasial kerapatan vegetasi di masa yang akan datang. Keterlibatan stakeholder dalam evaluasi dan umpan balik hasil prediksi adalah kunci untuk memastikan relevansi dan aplikasi praktis data remote sensing dalam berbagai konteks. Melalui komunikasi yang efektif, kolaborasi, dan penyesuaian model yang berkelanjutan, hasil prediksi dapat dioptimalkan untuk mendukung kebijakan dan keputusan berbasis data yang lebih baik.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Data

Data yang sudah diperoleh berupa gambar citra satelit yang diambil pada rentang tahun 2013 sampai 2024 dari USGS Landsat Collection 2 Level-2 lalu diolah menggunakan google colab. Dari data citra satelit tersebut lalu diolah untuk mendapatkan data NDVI pada kota Pangkalpinang. Normalized Difference Vegetation Index (NDVI) adalah indikator yang diperoleh dari citra satelit yang mengukur kondisi kesehatan tanaman [19]. NDVI

berhubungan langsung dengan kondisi kekeringan [20]. Selama musim hujan, NDVI lebih tinggi (maksimum +1) sedangkan pada musim kemarau NDVI turun (minimum -1) [21]. NDVI bersifat musiman tetapi karena perubahan iklim, perubahan naik/turun juga diamati selama bertahun-tahun [22]. Mengetahui tren dan siklus musiman NDVI terlebih dahulu dapat membantu perencanaan dan penerapan sumber daya yang lebih baik bagi pemangku kepentingan lokal dan nasional. Ada banyak parameter yang diketahui dan tidak diketahui yang memengaruhi NDVI. Pendekatan pembelajaran mesin, khususnya jaringan saraf dengan LSTM, telah menjadi populer untuk memodelkan data sekuensial dan semakin banyak diterapkan dalam tugas deret waktu seperti forecasting [23]. Dari sekian banyak pendekatan, salah satunya adalah melatih model pembelajaran mesin dengan data masa lalu untuk mendapatkan prediksi [24]. Pada bagian pertama, NDVI diekstraksi dari data citra Landsat 8 deret waktu di platform Google Earth Engine (GEE) melalui Python API. Pada bagian berikutnya, model prediktif dibangun untuk memperkirakan NDVI menggunakan ARIMA. Data mentah diambil pada google earth engine dan dilakukan masking untuk mendapatkan titik koordinat yang sesuai lokasi kota Pangkalpinang.



**Gambar 2.** Data mentah citra satelit

Sumber kode untuk mendapatkan data mentah yang dimana data citra satelit masih banyak tertutupi oleh awan. Setelah itu dilakukan setup awal untuk analisis data geospasial dan time series menggunakan Google Earth Engine bersama pustaka lain yang relevan seperti pandas, numpy, matplotlib, dan pmdarima. Persiapan ini meliputi instalasi pustaka, autentikasi, dan inisialisasi, serta pengimporan alat-alat yang dibutuhkan untuk analisis lebih lanjut. Landsat 8 oli telah beroperasi sejak tahun 2013. Peneliti akan mengambil setiap gambar hingga saat ini dengan tutupan awan maksimum yang diizinkan sebesar 20%.



**Gambar 3.** Citra satelit area pangkalpinang

Citra yang sudah diproses ini menunjukkan daerah perkotaan besar dengan infrastruktur yang padat di tengah, dikelilingi oleh area hijau dan vegetasi. Ada juga fitur geografis yang penting seperti sungai besar dan area pesisir. Ini bisa menjadi bagian dari sebuah kota yang berkembang dengan interaksi antara wilayah perkotaan, alam, dan infrastruktur pesisir.

```
print('Total gambar yang didapatkan :', dataset_AOI.size().getInfo())  
Total gambar yang didapatkan : 73
```

**Gambar 4.** Sumber kode untuk mengecek jumlah gambar citra satelit

Pada rentang 11 tahun didapatkan total gambar citra satelit sejumlah 73 yang akan diproses untuk melakukan prediksi.

```
# Nama setiap Landsat 8 band
dataset_AOI.first().bandNames().getInfo()

['SR_B1',
 'SR_B2',
 'SR_B3',
 'SR_B4',
 'SR_B5',
 'SR_B6',
 'SR_B7',
 'SR_QA_AEROSOL',
 'ST_B10',
 'ST_ATRAN',
 'ST_CDIST',
 'ST_DRAD',
 'ST_EMIS',
 'ST_EMSD',
 'ST_QA',
 'ST_TRAD',
 'ST_URAD',
 'QA_PIXEL',
 'QA_RADSAT']
```

**Gambar 5.** Sumber kode untuk mengetahui Band yang tersedia

SR\_B5 adalah Band Near Infrared (NIR) SR\_B5 adalah salah satu band dalam koleksi citra Landsat 8 Surface Reflectance (SR), di mana "SR" berarti bahwa nilai-nilai reflektansi telah dikalibrasi untuk memperhitungkan atmosfer, sehingga merepresentasikan reflektansi permukaan sebenarnya. Near Infrared (NIR) sangat berguna dalam berbagai aplikasi penginderaan jauh, terutama untuk memantau vegetasi, karena vegetasi yang sehat sangat memantulkan cahaya pada panjang gelombang ini. SR\_B4, SR\_B3 dan SR\_B2 masing-masing adalah Band Merah, hijau dan Biru. Penting untuk memahami bahwa band-band ini berkontribusi pada analisis berbagai fenomena bumi, seperti monitoring vegetasi, kelembapan tanah, dan perubahan tutupan lahan.

## 3.2 Analisis Data

### 3.2.1 NDVI Analisis

NDVI setiap piksel gambar dalam gambar akan dihitung, Dimana, NIR adalah band SR\_B5 dan Red adalah band SR\_B4. Rata-rata NDVI akan dihitung pada setiap gambar untuk mendapatkan NDVI time-series dari AOI.

```
def addNDVI(image):
    ndvi = image.normalizedDifference(['SR_B5', 'SR_B4']).rename('NDVI')
    return image.addBands(ndvi)
with_ndvi = dataset_AOI.map(addNDVI)
```

**Gambar 6.** Sumber kode untuk mengetahui band yang tersedia

Sumber kode ini berfungsi untuk menambahkan nilai NDVI ke dalam setiap citra pada dataset dengan menghitung rasio antara band NIR dan RED. NDVI adalah indeks yang umum digunakan untuk menganalisis kesehatan vegetasi.

```
def meanNDVI(image):
    image = ee.Image(image)
    meanDict = image.reduceRegion(reducer = ee.Reducer.mean().setOutputs(['NDVI']),
    geometry = pangkalpinang_AOI,
    scale = image.projection().nominalScale().getInfo(),
    maxPixels = 100000,
    bestEffort = True);
    return meanDict.get('NDVI').getInfo()
```

**Gambar 7.** Sumber kode untuk menghitung rata-rata NDVI pada setiap gambar dalam koleksi

Sumber kode sebelumnya menghitung dan mengembalikan rata-rata nilai NDVI untuk wilayah pangkalpinang\_AOI menggunakan citra satelit di Google Earth Engine. NDVI rata-rata ini dapat digunakan untuk mengevaluasi kesehatan vegetasi di area tersebut.

```
listOfImages_ndvi = with_ndvi.select('NDVI').toList(with_ndvi.size())
ndvi_coll = []
for i in range(listOfImages_ndvi.length().getInfo()):
    image = ee.Image(listOfImages_ndvi.get(i-1))
    temp_ndvi = meanNDVI(image)
    ndvi_coll.append(temp_ndvi)
```

**Gambar 8.** Sumber kode untuk hitung rata-rata NDVI untuk setiap gambar

Sumber kode ini melakukan iterasi pada setiap citra dalam koleksi with\_ndvi yang telah memiliki band NDVI, menghitung rata-rata NDVI untuk setiap citra, dan menyimpannya dalam list ndvi\_coll. Keluaran dari kode ini adalah daftar nilai rata-rata NDVI dari seluruh citra dalam dataset yang telah diproses.

```
dates = np.array(with_ndvi.aggregate_array("system:time_start").getInfo())
day = [datetime.datetime.fromtimestamp(i/1000).strftime('%Y/%m/%d') for i in (dates)]
```

**Gambar 9.** Sumber kode untuk ekstrak 'tanggal' dari koleksi gambar

Sumber kode untuk mengambil data waktu pengambilan citra dari koleksi with\_ndvi, lalu memformat waktu tersebut menjadi string tanggal dengan format YYYY/MM/DD. Hasil akhirnya adalah daftar tanggal dari setiap citra dalam koleksi tersebut, yang akan digunakan untuk menganalisis perubahan temporal (berdasarkan waktu) dalam data NDVI.

```
ndvi_df = pd.DataFrame(ndvi_coll, index = day, columns = ['ndvi'])
ndvi_df.index = pd.to_datetime(ndvi_df.index, format="%Y/%m/%d")
ndvi_df.sort_index(ascending = True, inplace = True)
ndvi_df.tail(5)
```

**Gambar 10.** Sumber kode untuk buat dataframe dengan kolom 'hari' dan 'NDVI'

Sumber kode ini mengubah data NDVI dan waktu pengambilannya menjadi DataFrame Pandas, kemudian mengonversi indeksnya ke format tanggal, mengurutkannya berdasarkan waktu, dan menampilkan lima data NDVI terakhir. Kode ini berguna untuk analisis tren NDVI berdasarkan urutan waktu.

	ndvi
2023-08-28	0.448646
2023-08-28	0.449546
2023-09-13	0.447627
2023-09-13	0.445145
2023-12-18	0.487503

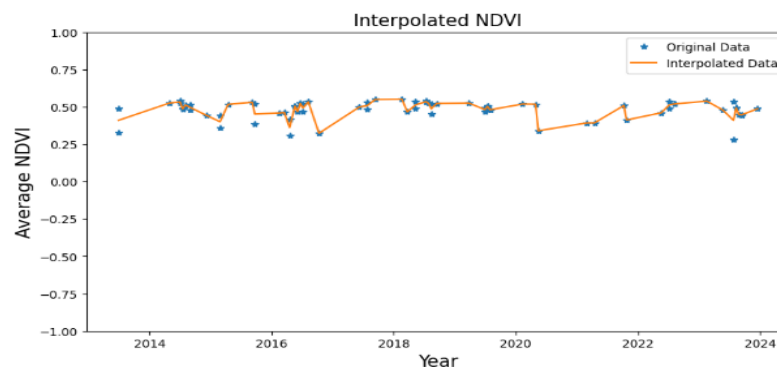
**Gambar 11.** Dataframe NDVI

Dataframe NDVI digunakan untuk memantau perubahan nilai NDVI pada beberapa tanggal di tahun 2023. Nilai NDVI yang lebih tinggi (mendekati 1) menunjukkan vegetasi yang lebih sehat atau lebat, sedangkan nilai yang lebih rendah (mendekati -1) menunjukkan sedikit atau tidak adanya vegetasi. Data ini dapat digunakan untuk analisis perubahan vegetasi di suatu wilayah pada rentang waktu tersebut.

```
plt.figure(figsize=(10,5), dpi=100)
plt.plot(ndvi_df, '*')
plt.plot(ndvi_df_daily)
plt.xlabel('Year', fontsize=15)
plt.ylabel('Average NDVI', fontsize=15)
plt.legend(['Original Data', 'Interpolated Data'])
plt.title("Interpolated NDVI", fontsize=15)
plt.ylim([-1, 1])
plt.show()
```

**Gambar 12.** Sumber kode untuk plotting interpolasi data ndvi

Sumber kode sebelumnya memvisualisasikan data NDVI asli dan hasil interpolasi dalam grafik. Grafik ini memiliki label sumbu, legenda, dan judul yang membantu memahami perbandingan antara data asli dengan data interpolasi. Nilai NDVI diplot dalam kisaran yang umum (-1 hingga 1).



**Gambar 13.** Interpolasi data NDVI

Data yang diinterpolasi selaras dengan data asli yang cukup untuk melatih algoritma untuk prediksi.

### 3.2.1. NDVI Prediksi Modeling

Ada banyak algoritma populer untuk pemodelan deret waktu, di antaranya ARIMA adalah model konvensional yang paling populer. Peneliti akan menggunakan ini untuk pemodelan prediksi NDVI.

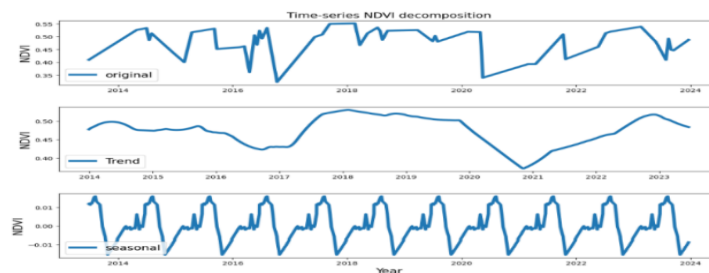
#### a. ARIMA Model

ARIMA merupakan singkatan dari Auto Regressive Integrated Moving Average. Model ARIMA secara umum dapat dinyatakan sebagai: prediksi  $Y(t) = \text{konstanta} + \text{lags kombinasi linier } Y(t) (\text{lags } p) + \text{kelambatan kombinasi linier kesalahan dalam istilah prediksi } (\text{lags } q)$ . Misalnya,  $Y(t) = 5 + 3 Y(t-1) + 2 \phi(t-1)$ . Di sini,

$p=q=1$ .  $d$  adalah singkatan dari diferensiasi yang dibutuhkan agar  $Y(t)$  menjadi seimbang. Ada banyak cara untuk menentukan kemungkinan kombinasi terbaik dari nilai  $p$ ,  $d$ , dan  $q$  untuk model ARIMA yang diberikan data deret waktu. Salah satu caranya adalah dengan memplot fungsi korelasi otomatis (ACF) dan fungsi autokorelasi parsial (PACF) secara manual untuk kombinasi parameter  $p$ ,  $d$ , dan  $q$  dan menentukan kasus untuk nilai ACF dan PACF mendekati nol. Peneliti akan menggunakan fungsi `auto_arma` yang akan menghasilkan parameter model yang dioptimalkan terbaik untuk rentang nilai  $p$ ,  $q$  berdasarkan Kriteria Informasi Akaike (AIC). AIC adalah penaksir yang menilai kualitas statistik suatu model. Model dengan nilai AIC yang lebih rendah menghasilkan kecocokan terbaik dengan data pelatihan dengan fitur paling sedikit. Fungsi `auto_arma` dapat melakukan pencarian grid pada parameter  $p$ ,  $d$ ,  $q$  (berkaitan dengan model ARIMA) dan  $P$ ,  $D$ , dan  $Q$  (berkaitan dengan komponen musiman), lalu melaporkan kembali model dengan nilai AIC terbaik.

b. Dekomposisi Musiman

Dapat dilihat komponen tren dan musiman dari data NDVI deret waktu. Pada Trend kecenderungan NDVI dalam kurun waktu tertentu. Misalnya jika NDVI meningkat/menurun dalam kurun waktu satu tahun. Di Seasonality merupakan variasi periodik dalam NDVI yang dapat dilihat setiap tahun. Ini menunjukkan bagian tahun mana NDVI meningkat/menurun dan itu terjadi secara siklus selama bertahun-tahun. Di residual merupakan komponen non sistematis dari nilai NDVI yang tidak terstruktur dan disebut sebagai noise.

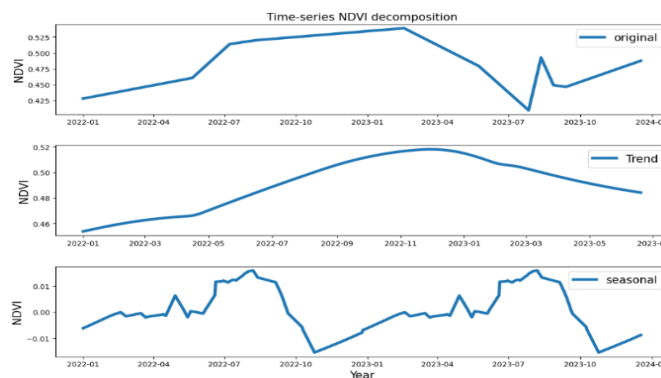


**Gambar 14.** Time-series NDVI Dekomposisi

Grafik tersebut menggambarkan data NDVI asli dari tahun 2013 hingga 2024. Grafik ini menunjukkan fluktuasi nilai NDVI selama periode tersebut, dengan beberapa puncak dan penurunan yang tajam. Nilai NDVI bervariasi antara sekitar 0,35 hingga 0,55, yang mencerminkan perubahan vegetasi dalam area studi selama waktu ini. Grafik kedua menunjukkan komponen tren dari data NDVI. Tren ini menunjukkan perubahan jangka panjang dalam NDVI, dengan garis yang lebih halus dibandingkan grafik asli. Dari grafik ini, bisa melihat bahwa ada fluktuasi besar dalam tren NDVI selama periode tersebut, dengan penurunan signifikan sekitar tahun 2021 dan peningkatan setelahnya. Grafik ketiga menunjukkan komponen musiman dari data NDVI. Komponen ini menangkap pola musiman yang berulang setiap tahun, dengan fluktuasi yang teratur. Grafik ini menunjukkan variasi tahunan dalam NDVI, yang kemungkinan besar mencerminkan perubahan vegetasi musiman seperti pertumbuhan dan penurunan vegetasi yang terjadi pada waktu-waktu tertentu dalam setahun. Secara keseluruhan, dekomposisi ini memungkinkan untuk memahami struktur data NDVI, memisahkan tren jangka panjang dari variabilitas musiman, dan memahami pola perubahan vegetasi dalam rentang waktu yang panjang.

c. NDVI Trend untuk 2 tahun

Dapat dilihat komponen tren dan musiman dari data NDVI deret waktu. Pada Trend kecenderungan NDVI dalam kurun waktu tertentu. Misalnya jika NDVI meningkat/menurun dalam kurun waktu dua tahun.



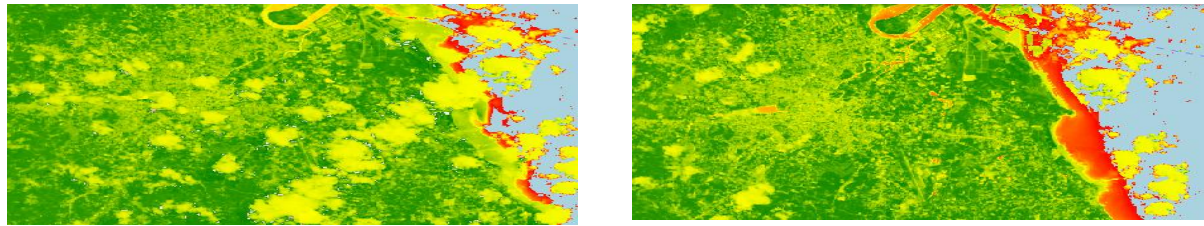
**Gambar 15.** Time-series NDVI Dekomposisi Musiman

Grafik tersebut menampilkan dekomposisi deret waktu dari tahun 2022 hingga awal tahun 2024, yang terbagi menjadi tiga komponen: original (asli), trend (tren), dan seasonal (musiman). Pada grafik pertama menunjukkan data NDVI asli dari tahun 2022 hingga Januari 2024. Grafik ini mengilustrasikan fluktuasi nilai

NDVI selama periode ini. Terlihat bahwa NDVI meningkat tajam dari awal 2022 hingga pertengahan 2022, mencapai puncak sekitar pertengahan hingga akhir 2022. Setelah itu, ada penurunan yang signifikan hingga pertengahan 2023, diikuti oleh fluktuasi dan sedikit kenaikan menjelang akhir tahun 2023. Grafik kedua menunjukkan komponen tren jangka panjang dari data NDVI. Tren ini memperlihatkan peningkatan yang konsisten dari awal 2022 hingga akhir 2022, diikuti oleh penurunan bertahap sepanjang tahun 2023. Grafik tren ini menunjukkan perubahan NDVI yang lebih halus dan berkelanjutan dibandingkan dengan data asli. Grafik ketiga menampilkan komponen musiman dari NDVI, yang mencerminkan pola musiman yang terjadi selama tahun 2022 hingga 2023. Komponen ini menunjukkan variasi NDVI yang berulang secara teratur setiap tahun, dengan pola kenaikan dan penurunan musiman yang konsisten. Ini menunjukkan pengaruh faktor musiman pada vegetasi, seperti perubahan musim tanam atau kondisi cuaca.

d. Gambar Peta dengan Nilai NDVI Rendah/Tinggi

Peneliti akan memilih satu hari sebagai sampel di bulan Mei dan Agustus 2022 dan memetakan peta NDVI untuk periode waktu ini.



(a). Mei

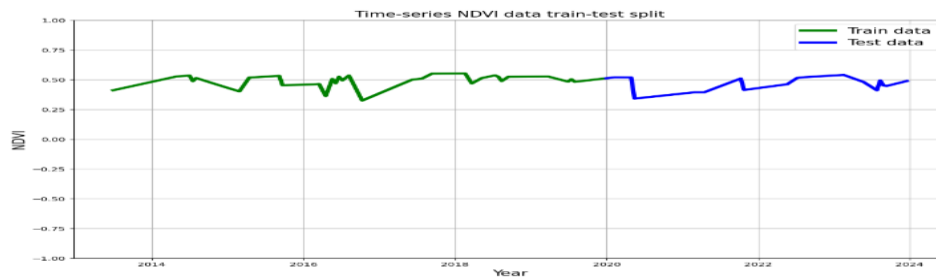
(b). Agustus

**Gambar 16.** Peta sebaran NDVI Pangkalpinang

Dengan skala yang sama, penyebaran warna kehijauan jauh lebih tinggi pada gambar bulan Mei dibandingkan gambar bulan Agustus.

e. Pemisahan Data Pelatihan dan Pengujian

Membagi data NDVI tujuh tahun pertama untuk rangkaian data latih dan empat tahun sisanya untuk data uji.



**Gambar 17.** Grafik pembagian data latih dan uji

Dari grafik time-series NDVI (Normalized Difference Vegetation Index) yang menampilkan pembagian data train dan test. Grafik menampilkan dua set data: data latih (train data) yang diwakili oleh garis hijau dan data uji (test data) yang diwakili oleh garis biru. Secara umum, nilai NDVI untuk data latih terlihat cukup stabil antara 0,4 dan 0,7 dari tahun 2014 hingga 2020. Data uji (setelah 2020) menunjukkan sedikit variasi yang lebih besar dibandingkan data latih, tetapi tetap berada di sekitar kisaran 0,4 hingga 0,5. Nilai NDVI yang stabil pada rentang 0,4 hingga 0,7 mengindikasikan bahwa vegetasi di area tersebut relatif konsisten, dengan fluktuasi kecil selama periode waktu yang diamati. Namun, terdapat beberapa penurunan singkat pada data latih di sekitar tahun 2016 dan 2020, yang mungkin mengindikasikan adanya gangguan vegetasi sementara misalnya karena perubahan cuaca, bencana alam, atau aktivitas manusia. Secara keseluruhan, grafik ini memberikan gambaran bahwa model pembelajaran mesin yang digunakan untuk memprediksi NDVI mampu menangkap pola yang cukup konsisten dari data historis.

f. Latih Model

Tentukan parameter terbaik untuk melatih model ARIMA dengan kumpulan data ini.

```
# fit_auto_arima model
model_autoARIMA = auto_arima(train_data, start_p=0, start_q=0, # nilai p, q minimum
                             test="adf", # menggunakan uji ADF untuk menentukan orde perbedaan 'd'
                             # dibutuhkan untuk penyeimbang
                             max_p=3, max_q=3, # nilai p dan q maksimum
                             m=7, # frekuensi seri, m=7 untuk pengamatan harian
                             d=1, # perbedaan satu langkah, d=1 dapat membuat data seimbang
                             # terlihat dari langkah pra-proses
                             seasonal=True, # musim yang diamati
                             start_p=0, # auto_arima akan kembali dengan parameter P, Q dan D yang optimal
                             D=0, # ketika parameter 'musiman' diaktifkan
                             trace=True,
                             error_action='ignore',
                             suppress_warnings=True,
                             stepwise=True) # step vs parallel. step direkomendasikan untuk kinerja yang lebih cepat
```

**Gambar 18.** Sumber kode latih model

Sumber kode ini bertujuan untuk mencari model ARIMA terbaik menggunakan `auto_arima` secara otomatis. Kode ini mempertimbangkan parameter musiman dan non-musiman, dengan melakukan pencarian parameter optimal berdasarkan nilai AIC (Akaike Information Criterion) melalui pendekatan stepwise. Parameter diferensiasi, musiman, dan batasan  $p$  serta  $q$  ditetapkan berdasarkan data latih yang digunakan.

```
Performing stepwise search to minimize aic
ARIMA(0,1,0)(0,0,1)[7] intercept : AIC=-25993.554, Time=3.07 sec
ARIMA(0,1,0)(0,0,0)[7] intercept : AIC=-25073.023, Time=0.45 sec
ARIMA(1,1,0)(1,0,0)[7] intercept : AIC=-29757.543, Time=8.95 sec
ARIMA(0,1,1)(0,0,1)[7] intercept : AIC=-24932.113, Time=4.19 sec
ARIMA(0,1,0)(0,0,0)[7] intercept : AIC=-25072.211, Time=0.19 sec
ARIMA(1,1,0)(0,0,0)[7] intercept : AIC=-30548.304, Time=0.78 sec
ARIMA(1,1,0)(0,0,1)[7] intercept : AIC=-28186.357, Time=2.58 sec
ARIMA(1,1,0)(1,0,1)[7] intercept : AIC=-29374.624, Time=2.99 sec
ARIMA(2,1,0)(0,0,0)[7] intercept : AIC=-30547.942, Time=1.91 sec
ARIMA(1,1,1)(0,0,0)[7] intercept : AIC=-30547.851, Time=2.97 sec
ARIMA(0,1,1)(0,0,0)[7] intercept : AIC=-26950.307, Time=3.07 sec
ARIMA(2,1,1)(0,0,0)[7] intercept : AIC=-30538.389, Time=5.20 sec
ARIMA(1,1,0)(0,0,0)[7] intercept : AIC=-30550.215, Time=0.44 sec
ARIMA(1,1,0)(1,0,0)[7] intercept : AIC=-29759.572, Time=3.26 sec
ARIMA(1,1,0)(0,0,1)[7] intercept : AIC=-28201.948, Time=5.06 sec
ARIMA(1,1,0)(1,0,1)[7] intercept : AIC=-29376.877, Time=2.82 sec
ARIMA(2,1,0)(0,0,0)[7] intercept : AIC=-30549.851, Time=1.49 sec
ARIMA(1,1,1)(0,0,0)[7] intercept : AIC=-30549.763, Time=1.04 sec
ARIMA(0,1,1)(0,0,0)[7] intercept : AIC=-27004.656, Time=0.92 sec
ARIMA(2,1,1)(0,0,0)[7] intercept : AIC=-30520.451, Time=2.66 sec

Best model: ARIMA(1,1,0)(0,0,0)[7]
Total fit time: 54.087 seconds
```

**Gambar 19.** Output best model

Proses stepwise search mengevaluasi berbagai kombinasi model ARIMA dengan tujuan memilih model terbaik berdasarkan nilai AIC yang terendah. Nilai AIC yang lebih rendah mengindikasikan model yang lebih baik karena AIC mempertimbangkan keseimbangan antara goodness of fit dan model complexity. Dari berbagai model yang diuji, model terbaik adalah ARIMA(1,1,0)(0,0,0)[7]. Model ini dipilih karena memiliki nilai AIC paling rendah yaitu 27004.656. Struktur model ARIMA(1,1,0) menunjukkan bahwa model menggunakan 1 autoregressive term (AR), 1 differentiation (I), dan 0 moving average term (MA). Di bagian musiman (seasonal), tidak ada komponen AR, I, atau MA yang digunakan, namun data diperlakukan dalam blok musiman 7 langkah (kemungkinan mingguan, jika data berbasis harian). Total waktu yang dibutuhkan untuk memilih model terbaik adalah 54.087 detik, dengan setiap percobaan model membutuhkan waktu eksekusi yang bervariasi, mulai dari beberapa detik hingga yang terlama sekitar 9 detik. Model dengan nilai AIC lebih rendah dipandang lebih efisien dalam memprediksi data dengan kompleksitas minimal. Nilai AIC bervariasi secara signifikan dari model ke model, dengan yang tertinggi mencapai lebih dari 39.000, dan model terbaik menghasilkan AIC sekitar 27.000. Secara keseluruhan model ARIMA(1,1,0)(0,0,0)[7] adalah pilihan terbaik berdasarkan hasil pencarian stepwise dengan tujuan meminimalkan AIC, yang berarti model ini diprediksi paling baik dalam menyeimbangkan antara akurasi prediksi dan kesederhanaan model.

```
import pandas as pd
import numpy as np
from statsmodels.tsa.arima.model import ARIMA
from sklearn.metrics import mean_squared_error
import itertools

# Asumsikan 'train_data' adalah data time series yang sudah disiapkan
# p, d, q adalah parameter ARIMA yang harus dioptimalkan
# train_data seharusnya adalah satu kolom pandas series atau numpy array

# Tentukan parameter p, d, q untuk ARIMA
p = range(0, 4) # bisa disesuaikan
d = range(0, 3) # bisa disesuaikan
q = range(0, 4) # bisa disesuaikan

# Menggunakan grid search untuk menemukan parameter terbaik
pdq = list(itertools.product(p, d, q))

best_aic = np.inf # Untuk menyimpan AIC terbaik
best_order = None
best_model = None

for param in pdq:
    try:
        model = ARIMA(train_data, order=param)
        fitted_model = model.fit()
        aic = fitted_model.aic
        if aic < best_aic:
            best_aic = aic
            best_order = param
            best_model = fitted_model
    except Exception as e:
        continue

# Cetak parameter terbaik
print(f"Best ARIMA order: {best_order}")
print(f"Best AIC: {best_aic}")

# Cetak ringkasan model terbaik
print(best_model.summary())
```

**Gambar 20.** Sumber kode bangun model arima dengan parameter terbaik

Sumber kode sebelumnya berfungsi untuk melakukan pencarian parameter terbaik untuk model ARIMA secara grid search. Tujuannya adalah menemukan kombinasi optimal dari parameter  $p$ ,  $d$ , dan  $q$  yang menghasilkan model dengan nilai AIC terendah, yang menyeimbangkan antara akurasi prediksi dan kompleksitas model. Grid search dapat memakan waktu yang lama, terutama jika rentang parameter yang diuji cukup besar. Penggunaan try-except membantu memastikan bahwa proses pencarian tetap berlanjut meskipun terjadi error pada beberapa kombinasi parameter.

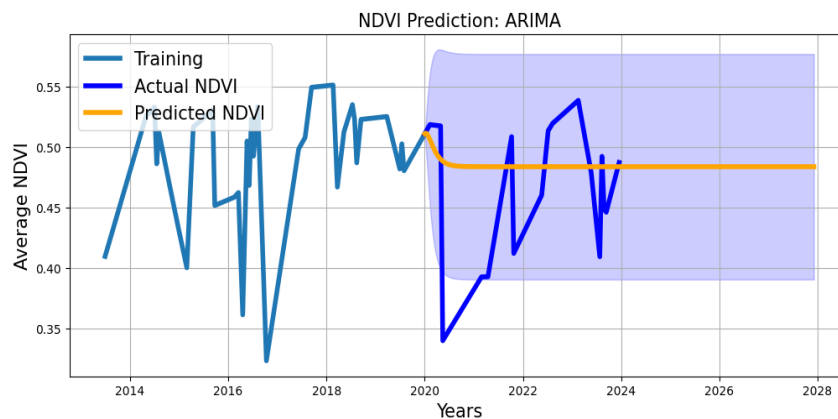
```

Best ARIMA order: (2, 0, 0)
Best AIC: -30563.697446350005
SARIMAX Results
-----
Dep. Variable:          ndvi      No. Observations:      2378
Model:                 ARIMA(2, 0, 0)  Log Likelihood         15285.849
Date:                  Thu, 26 Sep 2024  AIC                    -30563.697
Time:                  02:01:49        BIC                    -30540.601
Sample:                06-29-2013      HQIC                   -30555.291
                    - 01-01-2020
Covariance Type:      opg
-----
              coef      std err      z      P>|z|      [0.025      0.975]
-----
const          0.4840      0.012      39.025      0.000      0.460      0.508
ar.L1          1.9481      0.006     312.735      0.000      1.936      1.960
ar.L2         -0.9488      0.006    -153.194      0.000     -0.961     -0.937
sigma2         1.52e-07     1.28e-09     118.840      0.000     1.5e-07     1.55e-07
-----
Ljung-Box (L1) (Q):          0.86      Jarque-Bera (JB):      2767095.02
Prob(Q):                    0.35      Prob(JB):              0.00
Heteroskedasticity (H):     0.99      Skew:                  1.46
Prob(H) (two-sided):        0.88      Kurtosis:              170.09
-----

```

**Gambar 21.** Output latihan model terbaik

Model terbaik yang dipilih adalah ARIMA(2, 0, 0), yang berarti model menggunakan dua parameter autoregressive (AR), tidak ada differencing (d=0), dan tidak ada komponen moving average (MA). Nilai AIC (Akaike Information Criterion) terbaik yang diperoleh adalah -30563.697, yang mengindikasikan bahwa model ini memiliki performa yang baik dalam menyeimbangkan akurasi dan kompleksitas. Parameter AR(1) dan AR(2) signifikan dan menunjukkan hubungan yang kuat dengan variabel NDVI, yang artinya model autoregressive ini mampu menangkap pola data masa lalu dengan baik. Meskipun residual tidak berdistribusi normal (ditunjukkan oleh uji Jarque-Bera), tidak ada autokorelasi yang signifikan dalam residual dan model cenderung homoskedastik.

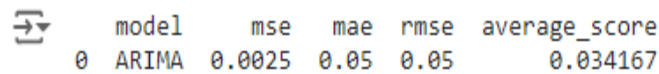


**Gambar 22.** Hasil prediksi menggunakan arima

Grafik tersebut menunjukkan prediksi NDVI menggunakan model ARIMA. Data pelatihan ditampilkan dengan garis biru muda, data pelatihan mencakup periode dari sekitar tahun 2013 hingga sekitar tahun 2020. Selama periode ini, nilai NDVI menunjukkan fluktuasi yang cukup signifikan. Data Aktual NDVI ditampilkan dengan garis biru tua, data aktual tampak dari sekitar tahun 2020 hingga 2024. Data ini menunjukkan fluktuasi yang lebih besar dibandingkan dengan data pelatihan, dengan beberapa penurunan tajam di awal 2021. Prediksi NDVI dengan garis oranye menunjukkan prediksi model ARIMA dari sekitar tahun 2020 ke depan. Prediksi ini tampak lebih stabil dibandingkan dengan data aktual dan tidak mengalami fluktuasi besar. Interval Kepercayaan dengan area berwarna biru muda di sekitar garis prediksi menggambarkan interval kepercayaan. Ini menunjukkan rentang ketidakpastian prediksi. Secara keseluruhan, grafik ini menunjukkan bahwa model ARIMA memprediksi stabilitas dalam kerapatan vegetasi di masa depan untuk Kota Pangkalpinang, namun terdapat ketidakpastian yang semakin besar pada tahun-tahun lebih jauh ke depan. ARIMA secara eksplisit dirancang untuk menangkap pola linier dalam data. Asumsinya adalah bahwa hubungan antara nilai masa lalu (lag) dan nilai saat ini dapat direpresentasikan dengan kombinasi linier. Jika pola dalam data bersifat non-linier, seperti adanya interaksi kompleks antar variabel atau perubahan mendadak, ARIMA tidak dapat menangkapnya dengan baik.

g. Performa model

Performa model dievaluasi dengan membandingkan nilai prediksi dengan data uji. Perbandingan akan dilakukan dalam bentuk Mean Squared Error (MSE), Mean Absolute Error (MAE), dan Root Mean Squared Error (RMSE). Rata-rata metrik ini akan diambil dan diberi nama 'average\_score' untuk mendapatkan gambaran kecocokan prediksi secara keseluruhan dengan data uji.



	model	mse	mae	rmse	average_score
0	ARIMA	0.0025	0.05	0.05	0.034167

**Gambar 23.** Output performa model

Evaluasi kinerja model ARIMA menghasilkan metrik MSE untuk model ARIMA adalah 0,0025. Metrik ini mengukur rata-rata perbedaan kuadrat antara nilai aktual dan nilai prediksi. MSE yang lebih rendah menunjukkan kinerja model yang lebih baik, karena mencerminkan lebih sedikit kesalahan besar. MAE adalah 0,05. Ukuran ini menunjukkan rata-rata perbedaan absolut antara nilai prediksi dan nilai aktual. Ini memberikan interpretasi langsung dari kesalahan prediksi rata-rata, dengan MAE yang lebih kecil menandakan akurasi prediksi yang lebih baik. RMSE adalah 0,05. Metrik ini, yang merupakan akar kuadrat dari MSE, juga menunjukkan besarnya rata-rata kesalahan prediksi. Karena RMSE menghukum kesalahan yang lebih besar lebih dari yang lebih kecil, RMSE yang lebih rendah menunjukkan bahwa model tersebut berkinerja baik dalam meminimalkan perbedaan besar. Skor rata-rata, dihitung sebagai rata-rata MSE, MAE, dan RMSE, kira-kira sebesar 0,0342. Metrik gabungan ini memberikan gambaran umum tentang kinerja model di berbagai ukuran kesalahan, yang mencerminkan tingkat akurasi dan keandalan secara keseluruhan. Singkatnya, model ARIMA menunjukkan kinerja yang kuat dengan metrik kesalahan yang rendah, yang menunjukkan efektivitasnya dalam membuat prediksi yang akurat untuk kumpulan data yang diberikan.

## 4. KESIMPULAN

Penelitian ini menunjukkan bahwa penggunaan algoritma machine learning berbasis time series, seperti ARIMA mampu menghasilkan model prediksi yang akurat dalam memproyeksikan kerapatan vegetasi di Kota Pangkalpinang. Root Mean Squared Error (RMSE): RMSE adalah 0,05. Metrik ini, yang merupakan akar kuadrat dari MSE, juga menunjukkan besarnya rata-rata kesalahan prediksi. Karena RMSE menghukum kesalahan yang lebih besar lebih dari yang lebih kecil, RMSE yang lebih rendah menunjukkan bahwa model tersebut berkinerja baik dalam meminimalkan perbedaan besar. Model ini memberikan wawasan yang bermanfaat dalam memantau perubahan vegetasi dari waktu ke waktu. Dengan menggunakan data NDVI dari citra satelit, model time series dapat mengevaluasi pola historis perubahan vegetasi dan memberikan prediksi kerapatan vegetasi di masa mendatang. Dengan dataset besar atau frekuensi tinggi, ARIMA bisa menjadi kurang efisien karena kompleksitasnya meningkat dengan panjang data.. Faktor spasial dan temporal memiliki peran penting dalam membentuk model prediksi. Dengan mempertimbangkan perubahan musiman dan tren jangka panjang, model yang dibangun mampu mengidentifikasi fluktuasi vegetasi yang bersifat musiman, sehingga prediksi yang dihasilkan dapat mendukung perencanaan tata kota yang lebih baik. Berdasarkan hasil penelitian, disarankan agar studi selanjutnya mengeksplorasi penggunaan metode machine learning lainnya, seperti Random Forest atau Long Short-Term Memory (LSTM), untuk membandingkan performa model dalam prediksi time series. Selain itu, integrasi dengan data eksogen seperti curah hujan, suhu, dan aktivitas manusia dapat meningkatkan ketepatan prediksi model dan memberikan wawasan yang lebih komprehensif terkait perubahan vegetasi. Hybrid models seperti ARIMA-LSTM atau ARIMA-CNN menawarkan pendekatan yang lebih holistik untuk menangkap pola linier, non-linier, dan spasial-temporal dalam data. Penerapan yang tepat dapat menghasilkan prediksi yang lebih akurat dan relevan untuk aplikasi di bidang remote sensing,

## REFERENCES

- [1] S. R. Shams, A. Jahani, S. Kalantary, M. Moeinaddini, and N. Khorasani, "The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO2 concentration," *Urban Clim.*, vol. 37, p. 100837, 2021.
- [2] Z. Zhou et al., "Analysis and prediction of vegetation dynamic changes in China: Past, present and future," *Ecol. Indic.*, vol. 117, p. 106642, Oct. 2020, doi: 10.1016/j.ecolind.2020.106642.
- [3] A. Y. Helmi, K. M. Harianto, and B. G. Dimmera, "Pemanfaatan Ruang Terbuka Hijau Dalam Meningkatkan Kualitas Lingkungan Kabupaten Bengkayang" *Jurdian*, 2023.
- [4] Z. Zhang, X. Wan, K. Sheng, H. Sun, L. Jia, and J. Peng, "Impact of Carbon Sequestration by Terrestrial Vegetation on Economic Growth: Evidence from Chinese County Satellite Data," *Sustainability*, vol. 15, no. 2, 2023, doi: 10.3390/su15021369.
- [5] Y. B. Pratama and N. P. Dalimunthe, "Implementasi Teknik Computer Vision Untuk Deteksi Viridiplantae Pada Lahan Pasca Tambang," *Bulletin of Computer Science Research*, vol. 3, no. 1, pp. 64–72, 2022, doi: 10.47065/bulletincsr.v3i1.193.
- [6] Q. Qin, D. Xu, L. Hou, B. Shen, and X. Xin, "Comparing vegetation indices from Sentinel-2 and Landsat 8 under different vegetation gradients based on a controlled grazing experiment," *Ecol. Indic.*, vol. 133, p. 108363, Dec. 2021, doi: 10.1016/J.ECOLIND.2021.108363.
- [7] A. Hardianto, P. U. Dewi, T. Feriansyah, N. F. S. Sari, N. S. Rifiana, and others, "Pemanfaatan Citra Landsat 8 Dalam Mengidentifikasi Nilai Indeks Kerapatan Vegetasi (NDVI) Tahun 2013 dan 2019 (Area Studi: Kota Bandar Lampung)," *J. Geosains Dan Remote Sens.*, vol. 2, no. 1, pp. 8–15, 2021.



- [8] I. K. Putrajaya and I. P. H. M. Martayana, “Analysis of Vegetation Density Using Landsat 8 Imagery for Mapping Characteristics of Green Open Space in Denpasar City,” *EAI*, 2021. doi: 10.4108/eai.10-11-2020.2303425.
- [9] Y. B. Pratama and A. Setiawan, “Implementasi Machine Learning Menggunakan Algoritma K-Means Untuk Klasifikasi Sekolah Dasar,” *Resolusi Rekayasa Tek. Inform. dan Inf.*, vol. 4, no. 3, pp. 249–257, 2024.
- [10] Y. A. Singgalen, “Tourism infrastructure development and transformation of vegetation index in Dodola Island of Morotai Island Regency,” *J. Inf. Syst. Informatics*, vol. 4, no. 1, pp. 130–144, 2022.
- [11] O. I. Gifari, K. Kusriani, and K. A. Yuana, “Analisis Perubahan Tutupan Lahan Menggunakan Metode Klasifikasi Terbimbing Pada Data Citra Penginderaan Jauh Kota Samarinda-Kalimantan Timur,” *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 18, no. 2, pp. 71–77, 2023.
- [12] D. Marlina, “Klasifikasi Tutupan Lahan pada Citra Sentinel-2 Kabupaten Kuningan dengan NDVI dan Algoritme Random Forest,” *STRING (Satuan Tulisan Ris. Dan Inov. Teknol.*, vol. 7, no. 1, pp. 41–49, 2022.
- [13] A. Y. Isnaeni and S. Y. J. Prasetyo, “Klasifikasi Wilayah Potensi Risiko Kerusakan Lahan Akibat Bencana Tsunami Menggunakan Machine Learning,” *J. Tek. Inform. Dan Sist. Inf.*, vol. 8, no. 1, pp. 33–42, 2022.
- [14] F. Ramadhani, R. Pullanagari, G. Kereszturi, and J. Procter, “Mapping of rice growth phases and bare land using Landsat-8 OLI with machine learning,” *Int. J. Remote Sens.*, vol. 41, no. 21, pp. 8428–8452, 2020.
- [15] A. S. Stepanov, E. A. Fomina, L. V. Illarionova, K. N. Dubrovin, and D. V. Fedoseev, “Vegetation indices (NDVI and EVI) time series approximation for monitoring crops of Khabarovsk territory,” *Informatics Autom.*, vol. 22, no. 6, pp. 1473–1498, 2023.
- [16] S. T. C. Wirawan and H. P. Chernovita, “Analysis of Vegetation Changes Using Satellite Imagery and Normalized Difference Vegetation Index (NDVI): A Case Study in Tuntang District, Semarang District,” *J. Inf. Syst. Informatics*, vol. 5, no. 4, pp. 1809–1820, 2023.
- [17] B. Haq et al., “Tech-Driven Forest Conservation: Combating Deforestation With Internet of Things, Artificial Intelligence, and Remote Sensing,” *IEEE Internet Things J.*, 2024.
- [18] R. Kumar, P. Kumar, and Y. Kumar, “Multi-step time series analysis and forecasting strategy using ARIMA and evolutionary algorithms,” *Int. J. Inf. Technol.*, vol. 14, no. 1, pp. 359–373, 2022.
- [19] H. Rakuasa and D. A. Sihasale, “Analysis of Vegetation Index in Ambon City Using Sentinel-2 Satellite Image Data with Normalized Difference Vegetation Index (NDVI) Method based on Google Earth Engine,” *J. Innov. Inf. Technol. Appl.*, vol. 5, no. 1, pp. 74–82, 2023.
- [20] S. M. Al-Hedny and A. S. Muhaimed, “Drought monitoring for Northern Part of Iraq using temporal NDVI and rainfall indices,” *Environ. Remote Sens. GIS Iraq*, pp. 301–331, 2020.
- [21] M. G. Ghebrezgabher, T. Yang, X. Yang, and T. E. Sereke, “Assessment of NDVI variations in responses to climate change in the Horn of Africa,” *Egypt. J. Remote Sens. Sp. Sci.*, vol. 23, no. 3, pp. 249–261, 2020.
- [22] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, “A comprehensive survey on graph neural networks,” *IEEE Trans. neural networks Learn. Syst.*, vol. 32, no. 1, pp. 4–24, 2020.
- [23] H. Wang, J. Zhang, and J. Yang, “Time series forecasting of pedestrian-level urban air temperature by LSTM: Guidance for practitioners,” *Urban Clim.*, vol. 56, p. 102063, Jul. 2024, doi: 10.1016/J.UCLIM.2024.102063.
- [24] R. Ahmad, B. Yang, G. Ettlin, A. Berger, and P. Rodriguez-Bocca, “A machine-learning based ConvLSTM architecture for NDVI forecasting,” *Int. Trans. Oper. Res.*, vol. 30, no. 4, pp. 2025–2048, 2023.