

Analisis Algoritma *Random Forest* dan Kombinasi Indeks Spektral untuk Identifikasi Lahan Terbangun (Kasus Kota Surakarta)

Ardia Tiara Rahmi¹, Kholis Hapsari Pratiwi², Delista Putri Deni³

^{1,2,3}Sekolah Vokasi Universitas Sebelas Maret, Jawa Tengah, Indonesia

e-mail : ardia_tiara@staff.uns.ac.id², kholis.hapsari@staff.uns.ac.id³, delistaputri@staff.uns.ac.id³

Abstract

In an effort to realize one of the objectives of the Sustainable Development Goals (SDGs) program goal 11 concerning Sustainable Cities and Settlements, controlling the intensity of urban built-up land in Indonesia really needs attention. One effort to monitor the condition of changes in built-up land that is easy, fast, cheap and efficient is to use the results of remote sensing data processing. Technological developments in the field of remote sensing are currently leading to processing based on Big Data and cloud computing, one of which is GEE (Google Earth Engine). Google Earth Engine. Identification of built-up land was carried out using the guided classification method of Machine Learning *Random Forest* and a combination of Spectral Index algorithms consisting of a combination of Urban Index (UI), Normalized Difference Vegetation Index (NDVI), and Modified Normalized Difference Water Index (MNDWI) algorithms. The object segmentation process on the results of the combined spectral index is carried out using the Otsu thresholding method. The results showed that the built-up area from the results of identification using the *Random Forest* algorithm and the spectral index combination algorithm in 2019 to 2023 has increased the area of land, which is shown in the 2019-2023 built-up land map. The accuracy test using the confusion matrix showed that the results of identification using a combination of spectral indices obtained OA and Kappa values which were included in the medium to high category, namely 98.69% and 0.969. So that this method can then continue to be used to monitor the pattern of development of built-up land in Surakarta City.

Keywords : Google Earth Engine, Spectral Index, Built-up Land, *Random Forest*, Sentinel-2

Abstrak

Dalam upaya mewujudkan salah satu tujuan dari program Sustainable Development Goals (SDGs) tujuan 11 tentang Kota dan Pemukiman yang Berkelanjutan, pengendalian intensitas lahan terbangun perkotaan di Indonesia sangat perlu diperhatikan. Salah satu upaya pemantauan kondisi perubahan lahan terbangun yang mudah, cepat, murah dan efisien adalah menggunakan hasil pengolahan data penginderaan jauh. Perkembangan teknologi di bidang penginderaan jauh saat ini mengarah ke pengolahan berbasis Big Data dan cloud computing salah satunya adalah GEE (Google Earth Engine). Dalam penelitian ini, akan dilakukan model pemantauan lahan terbangun dari data Citra Satelit Sentinel 2A SR resolusi 10 meter multi temporal menggunakan Google Earth Engine. Identifikasi lahan terbangun dilakukan dengan metode klasifikasi terbimbing Machine Learning *Random Forest* serta kombinasi algoritma Indeks Spektral yang terdiri dari kombinasi algoritma Urban Index (UI), Normalized Difference Vegetation Index (NDVI), dan Modified Normalized Difference Water Index (MNDWI). Proses segmentasi objek pada hasil kombinasi indeks spektral dilakukan menggunakan metode Otsu thresholding. Hasil penelitian menunjukkan bahwa lahan terbangun dari hasil identifikasi menggunakan algoritma *Random Forest* dan algoritma kombinasi indeks spektral pada tahun 2019 ke 2023 mengalami peningkatan luasan lahan, yang mana ditunjukkan dalam peta lahan terbangun tahun 2019-2023. Uji akurasi menggunakan matriks konfusi menunjukkan bahwa hasil identifikasi menggunakan kombinasi indeks spektral mendapatkan nilai OA dan Kappa yang termasuk dalam kategori sedang hingga tinggi yaitu 98,69 % dan 0,969. Sehingga metode tersebut selanjutnya dapat terus digunakan untuk memantau pola perkembangan lahan terbangun di Kota Surakarta.

Kata kunci: Google Earth Engine, Indeks Spektral, Lahan Terbangun, *Random Forest*, Sentinel-2

1. PENDAHULUAN

Salah satu Kota padat penduduk di Indonesia adalah Kota Surakarta yang memiliki luas 46,72 km² dengan jumlah penduduk 578.906 jiwa per Desember 2021. Dengan demikian, kepadatan penduduk di kota ini mencapai 12.391 jiwa per km persegi [1]. Hal tersebut menuntut adanya kebijakan untuk mengendalikan dan mengarahkan pengembangan wilayah perkotaan sesuai Tujuan Pembangunan Berkelanjutan (*Sustainable Development Goals/* SGDs) yaitu Tujuan 11 tentang Kota dan Pemukiman Berkelanjutan yang ditetapkan oleh PBB 2015 silam. Dalam usaha mendukung kebijakan tersebut, maka dibutuhkan upaya pemantauan perubahan lahan terbangun yang mudah, cepat, dan efisien. Saat ini teknologi penginderaan jauh untuk pemantauan perubahan lahan terbangun lebih banyak digunakan dari pada survei terestrial karena identifikasi obyek dapat dilakukan secara cepat, efisien, multitemporal, mencakup wilayah yang cukup luas, dan validitasnya dapat diandalkan [2]. Permodelan data inderaja memerlukan pengolahan data citra satelit dengan volume yang sangat besar dan tidak terstruktur sehingga membutuhkan teknologi / sumber daya baru yang lebih efisien [3]. Hadirnya sumber daya *cloud computing platform* seperti *Google Earth Engine* (GEE), dapat menjawab permasalahan Geo Big Data tersebut [4]. Algoritma *Random Forest* (RF) dapat digunakan sebagai pendekatan baru untuk memetakan tutupan lahan melalui *Google Earth Engine* [5].

Dalam penelitian oleh Kurniawan, dkk. (2017) [6], pendeteksian lahan terbangun berhasil dilakukan dengan kombinasi indeks spektral *Urban Index* (UI), *Normalized Difference Built-up Index* (NDBI) dan *Normalized Difference Vegetation Index* (NDVI) pada citra Landsat 8 Tahun 2013 di Kota Magelang. Kombinasi indeks spektral UI – NDVI dapat digunakan untuk mendeteksi lahan terbangun dengan hasil terbaik dibanding kombinasi NDBI – NDVI, dengan memperhatikan objek kenampakan citra berupa lahan vegetasi. Metode ekstraksi lahan terbangun dengan Kombinasi Indeks Spektraleberapa indeks spektral yaitu NDBI, NDVI, *Modified Normalized Difference Water Index* (MNDWI), *Normalized Difference Water Index* (NDWI) dan *Soil Adjusted Vegetation Index* (SAVI) pada citra Landsat 8 Tahun 2015 di Kota Yogyakarta dilakukan oleh Hidayati, dkk. (2018) [7]. Penelitian ini berhasil dilakukan dengan hasil terbaik dalam mengidentifikasi lahan terbangun yang terbebas dari badan air pada kombinasi indeks spektral NDBI–MNDWI. Kemudian Penelitian yang mengkaji perbandingan kemampuan beberapa metode *machine learning* dalam GEE untuk pemetaan tutupan lahan telah dilakukan oleh Mustofa pada tahun 2018 [8]. Penelitian tersebut membandingkan metode klasifikasi berbasis piksel menggunakan *machine learning* diantaranya *Winnow*, *SVM*, *Random Forest*, *Perceptron*, *Naïve Bayes*, *Minimum Distance*, *GmoMaxEnt*, *Ikpamir*, dan *CART*. Hasilnya menunjukkan bahwa metode *Random Forest* menunjukkan performa terbaik dalam identifikasi obyek dengan nilai overall accuracy sebesar 88.1%, yang artinya hasil identifikasi obyek menggunakan metode tersebut dapat diandalkan validitasnya hingga 88.1%.

Berdasarkan beberapa penelitian terdahulu mengenai analisis lahan terbangun, penulis menemukan suatu kajian yang belum dilakukan yaitu bagaimana jika identifikasi lahan terbangun dilakukan dengan mengkombinasikan



beberapa indeks spektral yaitu UI, NDVI dan MNDWI dari data citra satelit resolusi tinggi Sentinel 2A SR untuk interpretasi objek di permukaan bumi, kemudian membandingkannya dengan hasil identifikasi lahan terbangun menggunakan metode algoritma *Random Forest*. Kegiatan-kegiatan pengolahan data penelitian akan membutuhkan waktu yang lama jika diolah menggunakan aplikasi pengolahan data berbasis desktop, terutama jika area studinya sangat luas. Hadirnya platform pengolahan geo-big data berbasis cloud bernama *Google Earth Engine* (GEE) bisa menjadi salah satu solusi dari permasalahan tersebut [9]. *Google Earth Engine* (GEE) menyediakan *platform cloud* berbahasa pemrograman *JavaScript* untuk mengakses dan memproses citra penginderaan jauh yang tersedia secara bebas dalam jumlah besar, sehingga dapat melakukan pemantauan perubahan objek permukaan bumi secara multi-temporal dan kontinyu. Kemudian hasil penelitian ini akan mendapatkan metode identifikasi lahan terbangun yang terbaik antara metode klasifikasi terbimbing menggunakan algoritma *Random Forest* dan metode klasifikasi menggunakan kombinasi indeks spektral. Perbandingan dilakukan dari hasil uji akurasi yang menghasilkan nilai Kappa dari kedua metode tersebut. Metode yang terbaik akan dijadikan rekomendasi untuk pemantauan lahan terbangun secara kontinyu di Kota Surakarta.

2. METODOLOGI PENELITIAN

2.1. Studi Literatur

Studi literatur yang dipelajari yaitu perkembangan ruang perkotaan, faktor pendorong perubahan lahan terbangun, rencana tata ruang wilayah, aplikasi penginderaan jauh untuk kajian perkotaan dan tata ruang, Pengolahan data dengan *Google Earth Engine*, Sentinel-2 level 2A, Segmentasi berbasis *Otsu thresholding*, klasifikasi citra terbimbing *Random Forest*, kombinasi indeks spektral dengan algoritma UI, NDVI dan MNDWI, serta penelitian-penelitian terdahulu.

2.2. Pengumpulan Data

Pengumpulan data dilakukan untuk mengumpulkan segala data pendukung dalam penelitian. Adapun data yang akan diolah dalam penelitian ini meliputi:

- a) Data citra Sentinel-2 Level-2A pada tahun 2019, hingga 2023 di Kota Surakarta yang diimport dari *Copernicus Open Access Hub* (<https://scihub.copernicus.eu/>) ke sistem penyimpanan cloud yang disediakan oleh platform *Google Earth Engine*.
- b) Data estimasi jumlah populasi penduduk per grid-sel 100 m x 100 m pada tahun 2015 yang berasal dari WorldPop (<https://www.worldpop.org/>) dan diakses melalui *Google Earth Engine*. dalam format .tif.
- c) Peta batas administrasi Kota Surakarta Skala 1:25.000 yang diambil dari *The Global Administrative Unit Layers* (GAUL) dan dikeluarkan oleh Organisasi Pangan dan Pertanian (*Food and Agriculture Organization/FAO*). Data tersebut kemudian diimport ke dalam platform penyimpanan cloud GEE.

2.3. Kombinasi Kanal

Pada tahap ini dilakukan kombinasi dari beberapa kanal citra untuk menghasilkan beberapa jenis tampilan citra dan indeks spektral. Tampilan citra yang dibutuhkan pada penelitian ini adalah Natural color dengan komposisi kanal B4 (Red), B3 (Green) dan B2 (Blue), Tampilan *Natural Color* menampilkan objek pada permukaan bumi sesuai dengan kondisi dan warna aslinya. Tampilan tersebut kemudian digunakan untuk memudahkan peneliti melakukan interpretasi obyek pada citra dan membuat sampel *Training Sample*. Selain itu, kombinasi kanal juga dilakukan untuk menghitung algoritma NDVI, MNDWI, dan UI. NDVI atau indeks vegetasi membutuhkan Kombinasi Indeks Spektral4 (red) dan B8 (NIR). MNDWI atau indeks badan air membutuhkan kombinasi kanal B3 (green) dan B11 (SWIR 1). UI membutuhkan kombinasi kanal B8 (NIR) dan B12 (SWIR 2). Berikut ini adalah formula-formula indeks spektral tersebut.

$$NDVI = (\rho_{NIR} - \rho_{RED}) / (\rho_{NIR} + \rho_{RED}) \quad (1)$$

$$MNDWI = (\rho_{GREEN} - \rho_{SWIR1}) / (\rho_{GREEN} + \rho_{SWIR1}) \quad (2)$$

$$UI = (\rho_{SWIR2} - \rho_{NIR}) / (\rho_{SWIR2} + \rho_{NIR}) \quad (3)$$

Berikut ini adalah Gambar-Gambar script dalam *Google Earth Engine (GEE)*

```
// Function to calculate NDVI
function calculateNDVI(image) {
  return image.normalizedDifference(['B8', 'B4']).rename('NDVI');
}
// Function to calculate MNDWI
function calculateMNDWI(image) {
  return image.normalizedDifference(['B3', 'B11']).rename('MNDWI');
}
// Function to calculate UI
function calculateUI(image) {
  return image.normalizedDifference(['B12', 'B8']).rename('UI');
}
```

Gambar 1. Skrip *Google Earth Engine* untuk perhitungan algoritma NDVI, MNDWI, dan UI

2.4. Penentuan *Training Sample*

Penentuan sampel *Training Sample* berfungsi untuk menentukan jenis-jenis obyek di permukaan bumi yang akan digunakan sebagai acuan dalam proses klasifikasi terbimbing menggunakan Algoritma *Random Forest*. Penentuan *Training Sample* harus memenuhi persyaratan akurasi, dengan jumlah minimal 3 (tiga) sampel setiap jenis objek. Penetapan *Training Sample* juga dapat dilakukan dengan menggunakan data acuan berupa peta dan pengenalan objek secara visual agar penentuan *Training Sample* dapat merepresentasikan objek penelitian secara menyeluruh dan sesuai dengan kondisi sebenarnya. Sampel objek penelitian akan terbagi dalam dua (2) kelas tutupan lahan yaitu: lahan terbangun (*built-up land*) dan non-lahan terbangun (*non-built-up*). Dapat dilihat pada Gambar berikut skrip untuk input *training point* kedalam analisis klasifikasi menggunakan algoritma *Random Forest*.



```
var training_point = non_builtup.merge(builtup);  
  
// Step 4: Make a training dataset by sampling regions.  
var training = klip5.select(bands).sampleRegions({  
  collection: training_point,  
  properties: ['built'],  
  scale: 20  
});  
  
// Step 7: Create a validation set by subsampling training set  
var random = training.randomColumn();  
var validation = random.filter(ee.Filter.gte('random', 0.7));  
var training2 = random.filter(ee.Filter.lt('random', 0.7));  
  
// Step 5: Create the classifier and classify the image  
var classifier = ee.Classifier.smileRandomForest(10).train({  
  features: training2,  
  classProperty: 'built',  
  inputProperties: bands  
});
```

Gambar 2. Skrip *Google Earth Engine* untuk analisis *Training sample* dalam menyusun *classifier* menggunakan algoritma *Random Forest*

2.5. Klasifikasi Terbimbing Algoritma *Random Forest*

Pada penelitian ini proses klasifikasi terbimbing (*supervised classification*) yang digunakan adalah metode klasifikasi *Random Forest* yang merupakan salah satu metode klasifikasi dengan mengacu pada *Training Sample* dan dianalisis secara *Machine Learning*. Penetapan *Training Sample* dilakukan dengan menggunakan data acuan berupa peta dan pengenalan objek secara visual agar penentuan *Training Sample* dapat merepresentasikan objek penelitian sesuai dengan kondisi sebenarnya. Objek penelitian akan terbagi dalam dua kelas tutupan lahan yaitu lahan terbangun dan non-lahan terbangun. Objek penelitian akan terbagi dalam dua (2) kelas tutupan lahan yaitu: lahan terbangun dan (2) non-lahan terbangun. Selanjutnya dilakukan pembagian kelas tutupan lahan dari hasil klasifikasi terbimbing dan mengatur ulang rentang nilai piksel citra yang masuk kedalam kelas-kelas tersebut untuk mempermudah interpretasi secara visual. Sama halnya dengan proses-proses pengolahan data citra sebelumnya, klasifikasi terbimbing ini dilakukan menggunakan *JavaScript* melalui *Google Earth Engine* (GEE).

2.6. Perhitungan Kombinasi Indeks Spektral

Perhitungan indeks spektral dilakukan dua tahap yaitu:

(1) Perhitungan masing-masing indeks spektral (UI, NDVI dan MNDWI) sesuai dengan persamaan masing-masing indeks spektral dan,

(2) Perhitungan kombinasi indeks spektral yaitu

$$\text{Kombinasi Indeks Spektral} = UI - NDVI - MNDW \quad (4)$$

Pemilihan jenis kombinasi ini bertujuan untuk menguji kemampuan UI dalam identifikasi lahan terbangun dengan memperhatikan kenampakan objek vegetasi (NDVI) dan badan air (MNDWI).

2.7. Pembuatan Peta Lahan Terbangun

Setelah didapatkan dua hasil kombinasi indeks spektral Kombinasi algoritma UI – NDVI – MNDWI serta hasil klasifikasi terbimbing *Random Forest*, maka dilakukan pembuatan Peta Lahan Terbangun pada tahun 2019, 2020, 2021, dan 2022.

2.8. Perbandingan Metode Identifikasi Lahan Terbangun dari Hasil Uji Akurasi

Pada tahap ini dilakukan penentuan dan membandingkan metode identifikasi lahan terbangun yang dapat digunakan secara kontinyu. Proses ini bertujuan untuk mengetahui metode yang tepat digunakan di wilayah perkotaan dengan melihat nilai OA dan Kappa yang dihasilkan dari Uji Akurasi metode indentifikasi. Uji akurasi dilakukan untuk mengetahui tingkat ketelitian dalam mengidentifikasi lahan terbangun. Cara yang dilakukan adalah membandingkan hasil identifikasi obyek dari metode kombinasi indeks spectra dan algoritma *Random Forest* dengan data referensi berupa hasil interpretasi citra yang didapatkan dari proses digitasi kenampakan lahan terbangun melalui tampilan Natural Color. Proses pengujian dilakukan dengan menggunakan perhitungan metode *confusion matrix*. Metode penentuan titik sampel yang digunakan adalah *stratified random sampling* yaitu menentukan beberapa titik sampel yang didistribusikan secara acak dalam setiap kelas, di mana setiap kelas memiliki sejumlah poin yang sebanding dengan area relatifnya. Hasil uji akurasi menggunakan *confusion matrix* dengan mempertimbangkan nilai OA (overall accuracy) $\geq 80\%$ (Congalton dan Green,2009). Gambar berikut menunjukkan skrip analisis uji akurasi hasil identifikasi lahan terbangun menggunakan klasifikas *Random Forest* dan kombinasi indeks spektral pada *Google Earth Engine*.

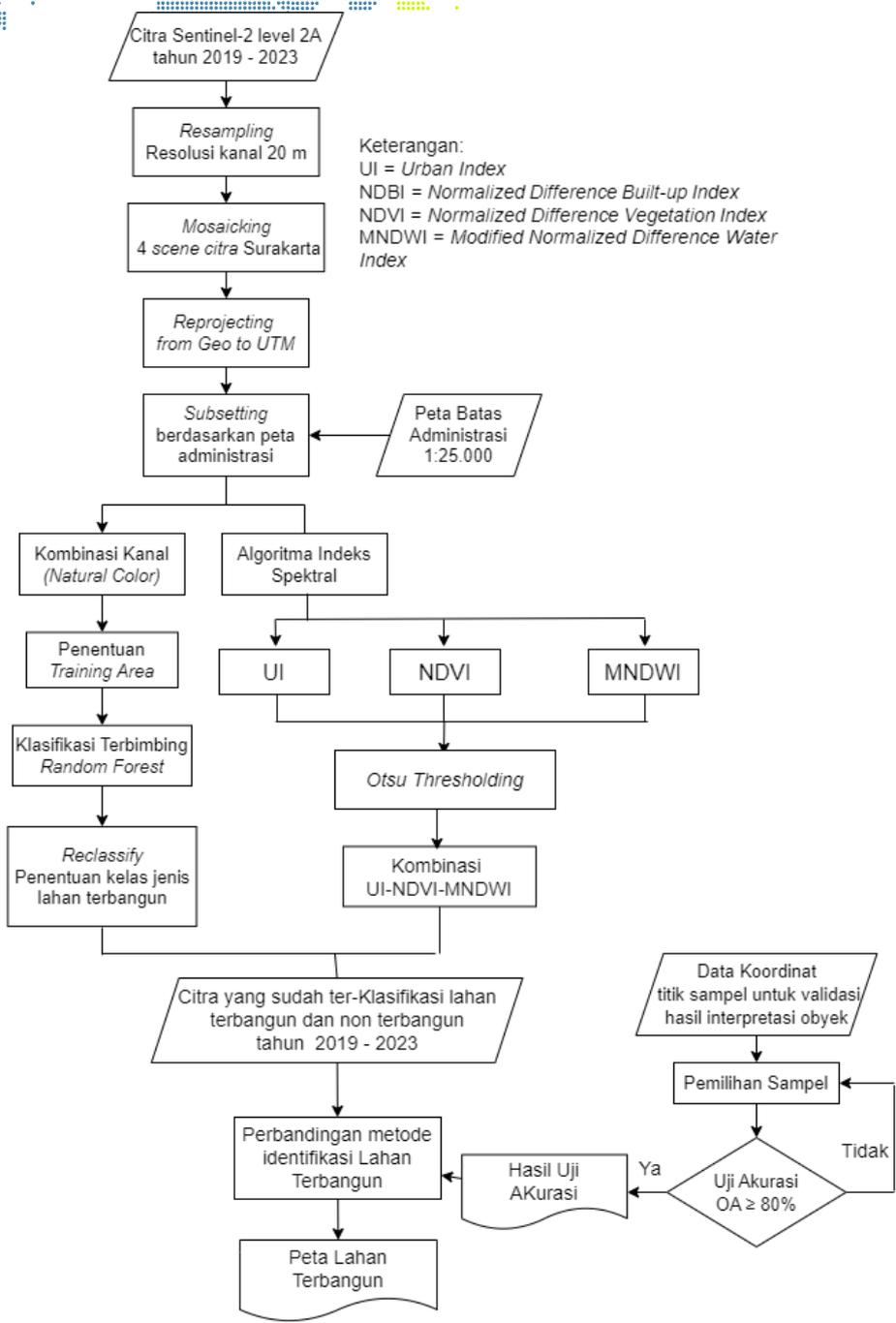
```
var trainAccuracy = classifier.confusionMatrix();
print('Resubstitution error matrix: ', trainAccuracy);
print('Training overall accuracy: ', trainAccuracy.accuracy());
print('Kappa: ', trainAccuracy.kappa());
print('User Accuracy: ', trainAccuracy.consumersAccuracy());
print('Producer Accuracy: ', trainAccuracy.producersAccuracy());

var validated = validation.classify(classifier);
// Get a confusion matrix representing expected accuracy.
var testAccuracy = validated.errorMatrix('built', 'classification');
print('Confusion matrix: ', testAccuracy);
print('Overall accuracy: ', testAccuracy.accuracy());
```

Gambar 3. Skrip Google Earth Engine untuk uji akurasi hasil klasifikasi lahan terbangun

2.9. Diagram Alir (Flow Chart)

Adapun diagram alir pengolahan data pada penelitian ini terlihat pada Gambar dengan penjelasan sebagai berikut :



Gambar 4. Diagram Alir Pengolahan Data

3. HASIL DAN PEMBAHASAN

Berikut ini hasil dan pembahasan dari proses pengolahan data Citra Sentinel 2A.

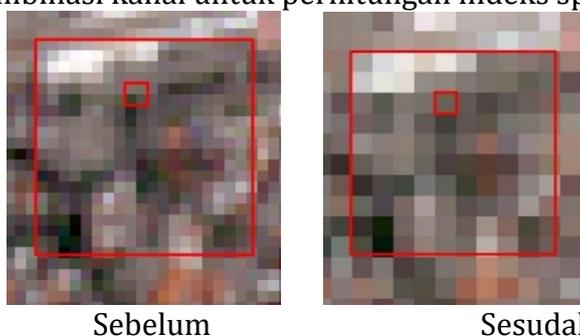
3.1. Hasil Pemrosesan Awal (*Preprocessing*) Sentinel-2

Pada pemrosesan awal data satelit dilakukan beberapa tahapan yaitu *resampling*, *mosaicking*, *reprojecting* dan *subsetting*. Seperti yang sudah dijelaskan pada subbab sebelumnya bahwa Citra Sentinel level 2A tidak perlu dikoreksi

atmosferik dan geometris. Dengan wilayah penelitian yang meliputi Kota Surakarta, maka dibutuhkan dua *scene* citra per tahunnya untuk mendapatkan wilayah penelitian yang utuh. Sedangkan kanal-kanal yang digunakan yaitu B2 (*Blue*), B3 (*Green*), B4 (*Red*), B8 (NIR), B11 (SWIR-1) dan B12 (SWIR-2).

3.2. Hasil *Resampling*

Pada tahap ini dilakukan proses *downscaling* atau menurunkan resolusi spasial kanal dari 10 m menjadi 20 m mengikuti skala spasial terendah dari kanal-kanal yang akan diolah. Kanal-kanal yang memiliki resolusi spasial 10 m diantaranya B2, B3, B4 dan B8, sedangkan kanal yang memiliki resolusi spasial asli 20 m adalah B11 dan B12. Pada proses *resampling* ini tidak disarankan melakukan *upscaling* untuk mencegah terjadinya kesalahan dalam interpolasi piksel. Penyamaan resolusi spasial di semua kanal merupakan syarat utama sebelum memasuki tahap kombinasi kanal untuk perhitungan indeks spektral.



Gambar 5. Perbandingan Resolusi Spasial Sebelum dan Sesudah *Resampling*

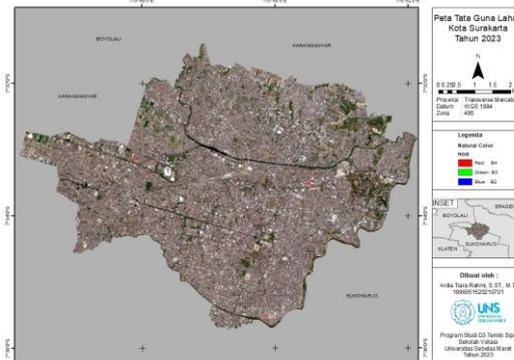
Berdasarkan Gambar 5, terlihat perbandingan sebelum dan sesudah proses *resampling* yang ditandai dengan berubahnya ukuran piksel (*cell size*) pada sumbu X dan Y dari 10 m pada kanal B2, B3, B4 dan B8 menjadi 20 meter. Dalam hal ini, untuk setiap piksel hasil *resampling* yang berukuran 20 meter pada sumbu X dan sumbu Y terbentuk dan terdiri dari 4 piksel yang berukuran 10 meter pada sumbu X dan sumbu Y. Tahap *resampling* dilakukan menggunakan analisis *JavaScript* melalui *Google Earth Engine*.

3.3. Hasil *Mosaicking* dan *Subsetting*

Dikarenakan pada wilayah penelitian membutuhkan dua *scene* citra Sentinel-2 pada setiap tahunnya maka dilakukan tahap *mosaicking* kedua *scene* citra tersebut menjadi satu *scene* yang utuh. Proses ini dilakukan pada tiap-tiap kanal yang digunakan dalam penelitian yaitu B2, B3, B4, B8, B11 dan B12. Untuk nilai piksel pada hasil *mosaicking* dilakukan akumulasi menggunakan rata-rata aritmatika dari semua piksel citra yang valid pada masing-masing *scene* citra yang dimasukkan dalam proses *mosaicking*.

Untuk mempermudah pemrosesan data citra Sentinel-2 khususnya pada wilayah penelitian dilakukan *subsetting* atau pemotongan bagian citra sesuai dengan batas administrasi Kota Surakarta. Proses ini juga dilakukan pada kanal-kanal citra yang digunakan di setiap tahunnya. Hasil dari *subsetting* adalah Gambar

dengan kanal B2, B3, B4, B8, B11 dan B12 yang hanya mencakup piksel di Kota Surakarta.

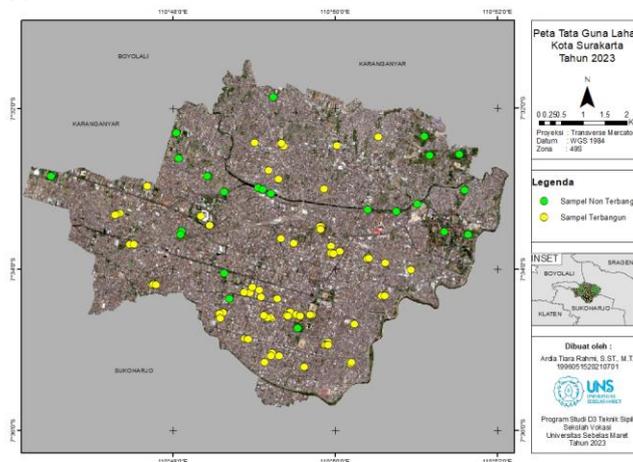


Gambar 6. Peta Kota Surakarta dengan Tampilan *Natural Color*.

Gambar di atas merupakan Peta Kota Surakarta dengan tampilan *Natural Color* yang dihasilkan dari kombinasi kanal RGB atau B4, B3, dan B2. Citra pada Gambar tersebut sudah melalui proses *resampling*, *mosaicking*, *reprojecting* dan *subsetting*.

3.4. Hasil Pengolahan Berbasis Klasifikasi *Random Forest*

Training sample yang digunakan dalam proses klasifikasi dibuat berdasarkan interpretasi visual pada citra yang berkomposit *natural color*. Jumlah kelas tutupan lahan yang dihasilkan ada dua yaitu: lahan terbangun (*built-up*) dan non-lahan terbangun (*non-built-up*), sehingga jumlah sampel *training sample* tutupan lahan untuk menghasilkan klasifikasi *Random Forest* juga menyesuaikan dengan jumlah kelas tutupan lahan tersebut. Bentuk *training sample* yang dipilih adalah titik koordinat agar mempermudah proses digitasi obyek yang lebih teliti. Penentuan *training sample* ini disebar sesuai dengan posisi lahan terbangun dan non-lahan terbangun yang terlihat pada citra *Natural Color* di seluruh wilayah penelitian seperti pada Gambar 7.

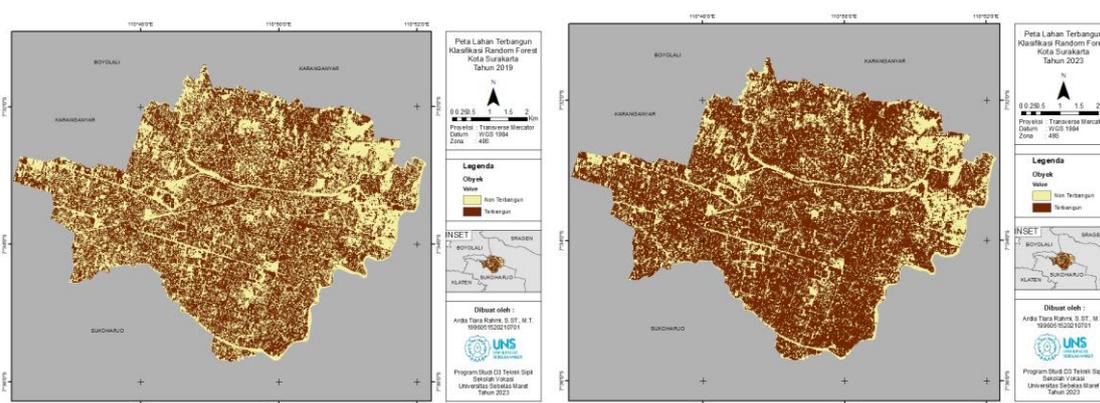


Gambar 7. Tampilan Sebaran Titik *Training Sample*

Tabel 1. Jumlah Training Sample Klasifikasi Terbimbing

Kelas Tutupan Lahan	Jumlah Training Sample	Interpretasi
Lahan Terbangun	105	●
Non Lahan Terbangun	105	●

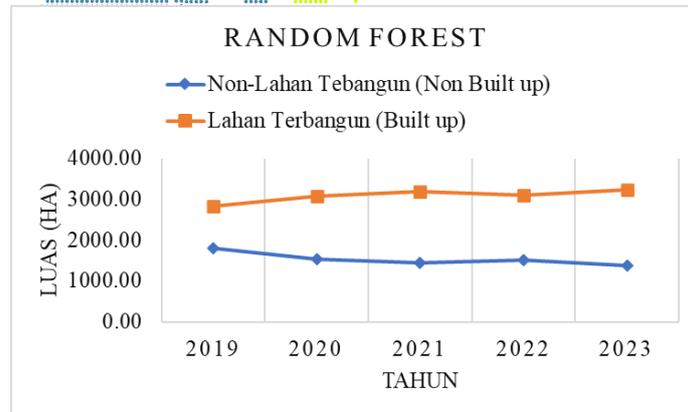
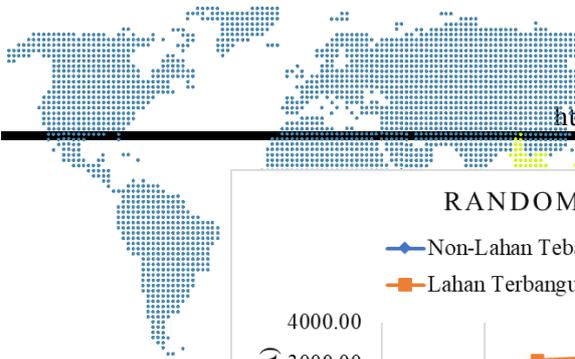
Penelitian ini menetapkan *training sample* berupa titik untuk digitasi objek lahan terbangun dan non lahan terbangun masing-masing sebanyak 100 titik. Obyek lahan terbangun meliputi piksel yang meliputi objek atap bangunan sedangkan objek non-lahan terbangun meliputi badan air dan vegetasi. Digitasi obyek pada penentuan *training sample* ini dilakukan pada semua citra dari tahun 2019 s/d 2023. Hal ini dilakukan untuk mencegah terjadinya kesalahan interpretasi akibat perubahan tutupan lahan yang terjadi dari tahun ke tahun. Hasil dari pembuatan Training Sample kemudian digunakan dalam proses klasifikasi citra Sentinel-2 *Random Forest*. Dari proses klasifikasi ini kemudian dihasilkan citra tutupan lahan terbangun Kota Surabaya tahun 2019 dan 2023 sebagaimana terlihat pada Gambar 8.



Gambar 8. Peta Tutupan Lahan hasil Klasifikasi *Random Forest*

Tabel 2. Perbandingan Luas Tutupan Lahan Hasil Klasifikasi *Random Forest* Tahun 2019 s/d 2023

Tutupan Lahan	2019 (ha)	2020 (ha)	2021 (ha)	2022 (ha)	2023 (ha)
<i>Non Built up</i>	1795,76	1544.72	1440,92	1523.50	1388.71
	38.81%	33.39%	31.14%	32.93%	30,02%
<i>Built Up</i>	2830,79	3081.84	3185.64	3103.05	3237.84
	61.19%	66.61%	68.86%	67.07%	69.98%
Total	4626.56	4626.56	4626.56	4626.56	4626.56

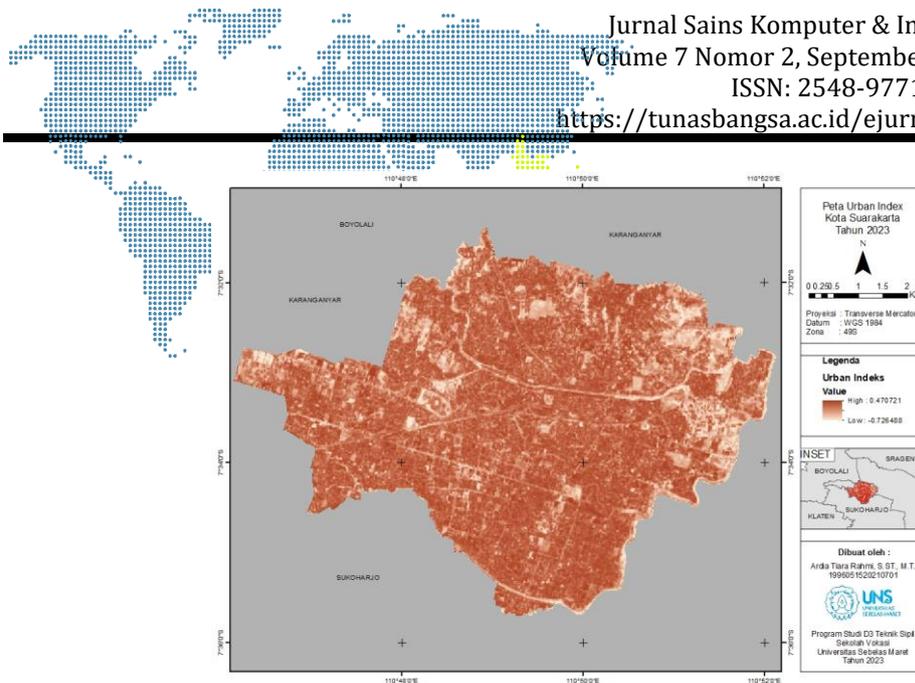


Gambar 9. Grafik Perbandingan Luas Tutupan Lahan Hasil Klasifikasi Random Forest Tahun 2019 s/d 2023

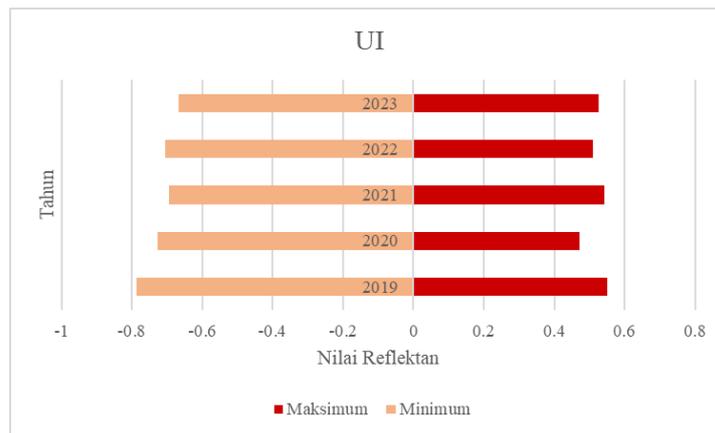
Berdasarkan Gambar 9, jika melihat grafik tersebut luas lahan terbangun dari tahun 2019 hingga tahun 2021 terus bertambah, namun pada tahun 2021 luas lahan terbangun menurun. Kemudian pada tahun 2023, luas lahan terbangun Kembali meningkat. Hal ini dapat dianalisis melalui jumlah piksel citra yang teridentifikasi sebagai lahan terbangun pada tahun 2022 lebih kecil dibandingkan pada tahun 2021. Hal ini diperkuat oleh informasi perhitungan luasan yang terlihat pada Tabel 2. Diketahui bahwa pada tahun 2023, kelas lahan terbangun memiliki luasan tertinggi sebesar 3237.84 Hektar (69,98 %). Sedangkan pada tahun 2019, kelas lahan terbangun memiliki luasan terendah sebesar 2830,79 Hektar (61.19%). Penurunan luas lahan terbangun dari tahun 2021 ke tahun 2022 dapat terjadi akibat beberapa faktor. Faktor pertama adalah kesalahan interpretasi objek tutupan lahan saat membuat *training sample*. Faktor lainnya adalah kualitas citra yang kurang baik pada tahun tersebut dimana tutupan awan sangat mempengaruhi proses identifikasi lahan. Jika melihat tampilan komposit *natural color* Kota Surakarta pada tahun 2022, ada beberapa wilayah yang tertutup awan dimana wilayah tersebut dapat menghasilkan nilai *null* atau piksel tak bernilai. Berdasarkan analisis penulis maka pada metode klasifikasi terbimbing ini masih terjadi misinterpretasi dalam mengidentifikasi objek lahan terbangun, oleh karena itu perlu dilakukan uji akurasi untuk dapat mengetahui seberapa besar misinterpretasi tersebut.

3.5. Hasil Analisis Algoritma Indeks Spektral UI

Indeks spektral yang diterapkan adalah UI untuk mengidentifikasi objek lahan terbangun, NDVI untuk mengidentifikasi objek vegetasi dan MNDWI untuk mengidentifikasi objek badan air. Pengolahan dilakukan dengan memasukkan rumus atau formula dari setiap indeks spektral yang digunakan dan dilakukan visualisasi dengan gradasi warna tertentu. Identifikasi objek lahan terbangun dengan indeks spektral UI (*Urban Index*) menggabungkan kemampuan dari kanal B8 (NIR) dan B12 (SWIR-12) yang diolah dengan formula (3). Adapun hasil visualisasi dari identifikasi objek lahan terbangun berbasis indeks UI terlihat pada Gambar 10.



Gambar 10. Peta Hasil Kombinasi Indeks Spektral UI

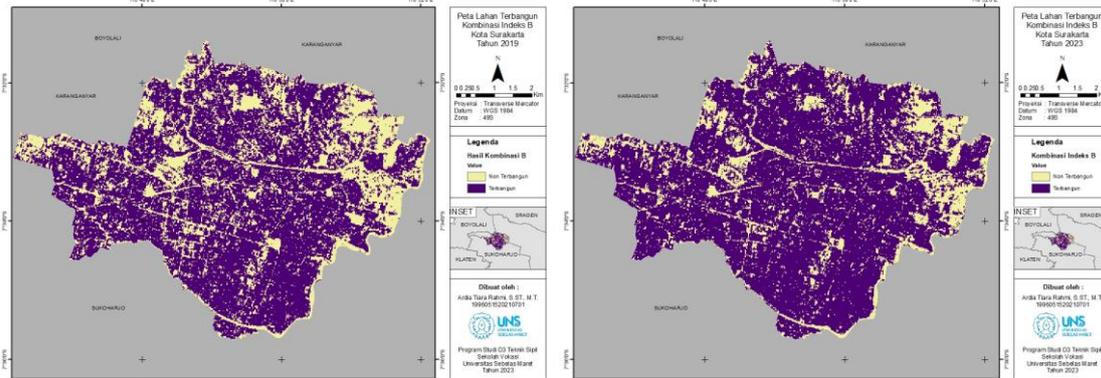


Gambar 11. Grafik Nilai Reflektan UI

Berdasarkan Gambar 11, semakin tinggi kemungkinan piksel tersebut diidentifikasi sebagai lahan terbangun, berlaku sebaliknya untuk non-lahan terbangun. Citra Sentinel yang digunakan pada proses analisis indeks spektral UI pada Gambar 10 tersebut adalah citra sentinel pada tahun 2023, tahun dengan kondisi citra yang paling relevan dengan kondisi saat ini.

3.6. Hasil Kombinasi Indeks Spektral

Hasil segmentasi obyek dari proses *Otsu thresholding* kemudian digunakan untuk pembentukan kombinasi indeks spektral. Pada penelitian ini dihasilkan kombinasi = UI - NDVI - MNDWI berdasarkan rumus nomor 4. Perhitungan yang dilakukan adalah melakukan pengurangan objek piksel dari masing-masing indeks spektral. Visualisasi hasil dari kombinasi indeks spektral ditampilkan pada Gambar 12.



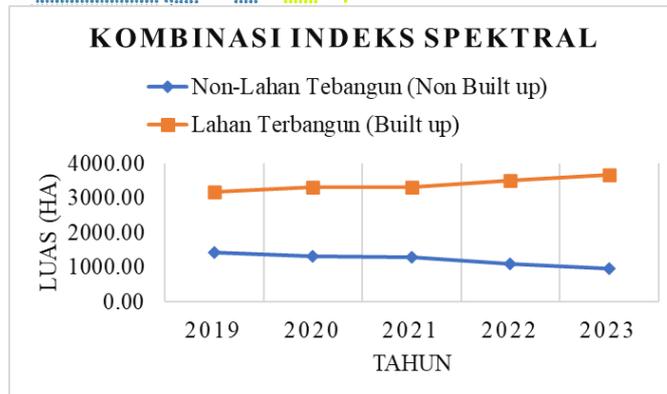
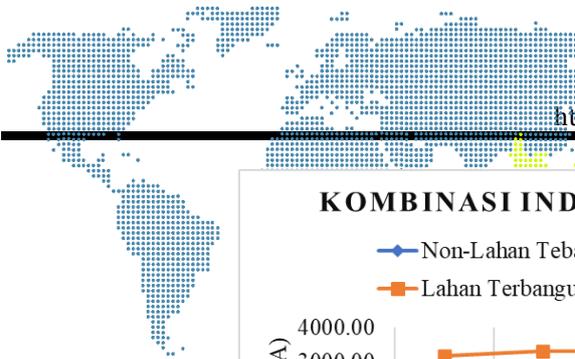
Gambar 12. Hasil Kombinasi Indeks Spektral tahun 2019 dan 2023

Diketahui bahwa kombinasi memiliki asumsi bahwa wilayah yang diidentifikasi sebagai objek lahan terbangun dari indeks UI dikurangkan dengan wilayah yang diidentifikasi sebagai objek vegetasi dari indeks NDVI dan wilayah yang diidentifikasi sebagai objek badan air dari indeks MNDWI. Berdasarkan Gambar 16, hasil yang didapatkan dari kombinasi ini adalah citra raster berupa piksel lahan terbangun (UI) yang sudah terbebas dari piksel objek vegetasi (NDVI) dan piksel objek badan air (MNDWI). Untuk kelas tutupan lahan terbagi menjadi dua yaitu lahan terbangun (*built-up*) yang memiliki nilai piksel sebesar 1 dan non-lahan terbangun (*non-built-up*) yang memiliki nilai piksel sebesar 0. Piksel 1 ditampilkan dengan warna gelap untuk menunjukkan perbedaan signifikan antara obyek lahan terbangun dan non lahan terbangun.

Tabel 3. Perbandingan Luas Tutupan Lahan Hasil Kombinasi tahun 2019 s/d 2023

Tutupan Lahan	2019	2020	2021	2022	2023
<i>Non Built up</i>	1438.71	1310,65	1292.00	1107.26	959.19
	31.10%	28.33%	27.93%	23.93%	20,73%
<i>Built up</i>	3187.84	3315.91	3334.56	3519.30	3667.37
	68.90%	71.67%	72.07%	76.07%	79.27%
Total	4626.56	4626.56	4626.56	4626.56	4626.56

Hasil perhitungan luas dari identifikasi lahan terbangun metode kombinasi indeks spektral menunjukkan pola yang berbanding lurus antara pertambahan tahun dan luas lahan terbangun. Semakin bertambah tahun makan jumlah luas lahan terbangun pun semakin meningkat. Namun hasil tersebut masih tetap harus diuji keakuratannya. Tabel 3 menunjukkan kelas tutupan lahan terbangun terbesar pada tahun 2023 yaitu 3667,37 Ha (79,27%), sedangkan kelas tutupan lahan terbangun terendah pada tahun 2019 yaitu sebesar 3187.84 Ha (68.90%). Untuk mengetahui perbandingan setiap tahun secara grafik dapat dilihat pada Gambar 13.



Gambar 13. Grafik Perbandingan Luas Hasil Kombinasi Indeks Spektral

Gambar 13 menunjukkan pola perkembangan lahan terbangun yang sama antara kombinasi A dan Kombinasi Indeks Spektral. Seiring bertambahnya tahun, luas lahan terbangun semakin meningkat. kedua hasil tersebut menunjukkan hasil yang baik dan searah dengan pola peningkatan kepadatan penduduk di Indonesia. Maka untuk mendapatkan metode identifikasi terbaik perlu dilakuka uji akurasi.

3.7. Hasil Uji Validasi Tutupan Lahan

Validasi tutupan lahan dilakukan terhadap piksel yang dihasilkan dari metode klasifikasi yang telah dilakukan yaitu berbasis klasifikasi terbimbing *Random Forest* dan klasifikasi berbasis kombinasi indeks spektral. Hal ini diterapkan dengan cara membandingkan data hasil klasifikasi terhadap data referensi yaitu Interpretasi melalui *Google Earth Pro* yang disesuaikan juga dengan citra sentinel *Natural Color*. Jumlah ukuran sampel dalam melakukan validasi tutupan lahan memiliki nilai yang sama dengan jumlah ukuran sampel dalam menentukan korelasi antar indeks spektral yang ditentukan berdasarkan persamaan distribusi multinomial.

3.8. Hasil Uji Validasi Klasifikasi *Random Forest*

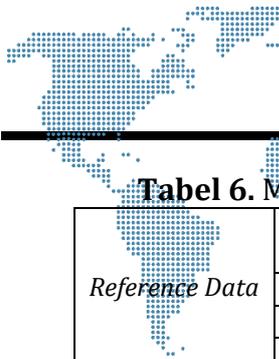
Berdasarkan hasil validasi tutupan lahan berbasis klasifikasi terbimbing dengan algoritma *Random Forest*, didapatkan nilai koefisien Kappa dengan nilai minimum sebesar 0,94 pada tahun 2019 dan nilai maksimum sebesar 0,98 pada tahun 2023. Sedangkan, nilai OA (*overal accuracy*) didapatkan nilai minimum sebesar 97,00% pada tahun 2019 dan nilai maksimum sebesar 99,42% pada tahun 2023.

Tabel 5. Matriks Konfusi Klasifikasi *Random Forest* Tahun 2019

Reference Data	Class	Classified Data			
		non-built-up	built up	Total	Producer Accuracy
	non-built-up	85	1	86	98,8372
	built up	4	77	81	95,0617
	Total	89	78	167	
	User Accuracy	95,5056	98,71794		

Total Sample: 167, Total Corrected Sample: 162

Overall Accuracy: 97,01% Kappa: 0,94



Tabel 6. Matriks Konfusi Klasifikasi *Random Forest* Tahun 2023

Reference Data	Class	Classified Data			Producer Accuracy
		non-built-up	built up	Total	
	non-built-up	84	0	84	100
	built up	1	86	87	98,8505
	Total	85	86	171	
	User Accuracy	98,8235	100		

Total Sample: 171, Total Corrected Sample: 170

Overall Accuracy: 99,415% Kappa: 0,988

Jika dilihat berdasarkan nilai Kappa dan OA tersebut dapat disimpulkan bahwa hasil klasifikasi terbimbing termasuk dalam kategori sangat dapat diandalkan yang berarti memiliki kekuatan kesepakatan pada tingkat tinggi. Adapun hasil perhitungan uji validasi dengan metode *confusion matrix* dapat dilihat pada Tabel 5 dan Tabel 6.

3.9. Hasil Uji Validasi Berbasis Kombinasi Indeks Spektral

Berdasarkan hasil validasi tutupan lahan berbasis kombinasi indeks spektral, didapatkan nilai koefisien Kappa dengan nilai minimum sebesar 0,933 pada tahun 2020 dan nilai maksimum sebesar 0,988 pada tahun 2021. Sedangkan, nilai OA (*overall accuracy*) didapatkan nilai minimum sebesar 96,67 % pada tahun 2020 dan nilai maksimum sebesar 99,41% pada tahun 2021.

Tabel 9. Matriks Konfusi Kombinasi Indeks Spektral Tahun 2020

Reference Data	Class	Classified Data			Producer Accuracy
		non-built-up	built up	Total	
	non-built-up	74	2	76	97,3684
	built up	3	71	74	95,9459
	Total	77	73	150	
	User Accuracy	96,1038	97,2602		

Total Sample: 150, Total Corrected Sample: 145

Overall Accuracy: 96,67% Kappa: 0,933

Tabel 10. Matriks Konfusi Kombinasi Indeks Spektral Tahun 2021

Reference Data	Class	Classified Data			Producer Accuracy
		non-built-up	built up	Total	
	non-built-up	82	1	83	98,7951
	built up	0	87	87	100
	Total	82	88	170	
	User Accuracy	100	98,8636		

Total Sample: 170, Total Corrected Sample: 169

Overall Accuracy: 99,39% Kappa: 0,988

3.10. Analisis Hasil Uji Validasi Keseluruhan

Berdasarkan hasil uji validasi yang telah dilakukan, maka pada penelitian ini dapat dirangkum bahwa nilai Kappa dan OA pada metode klasifikasi terbimbing termasuk dalam kategori sedang atau memiliki kekuatan kesepakatan pada tingkat

sedang, sedangkan pada metode kombinasi indeks spektral termasuk dalam kategori sedang hingga tinggi atau memiliki kekuatan kesepakatan pada tingkat sedang hingga tinggi. Seperti yang terlihat pada Tabel 11, berikut rangkuman nilai OA dan Kappa dari semua metode yang telah dilakukan.

Tabel 11. Nilai OA dan Kappa Hasil Penelitian Secara Keseluruhan

Tahun	Parameter	Metode	
		Kombinasi Indeks Spektral	Klasifikasi RF
2019	OA (%)	96,748	97,006
	Kappa	0,975	0,940
2020	OA (%)	96,667	97,024
	Kappa	0,933	0,940
2021	OA (%)	99,412	97,590
	Kappa	0,988	0,952
2022	OA (%)	99,306	97,576
	Kappa	0,986	0,951
2023	OA (%)	98,101	99,415
	Kappa	0,962	0,988
Rata-Rata	OA	98,047	97,722
	Kappa	0,969	0,954

Jika memperhatikan nilai OA dari setiap metode di tahun 2019 hingga 2023, Kombinasi Indeks Spektral memberikan nilai tertinggi, disusul dengan nilai OA dari hasil klasifikasi terbimbing *Random Forest*. Kedua metode tersebut memiliki nilai Kappa minimum yang masih termasuk dalam kategori tinggi, maka dapat dikatakan bahwa hasil pengolahan kedua metode tersebut dapat diterima dan sangat dapat diandalkan untuk merepresentasikan kondisi objek yang sebenarnya.

Hasil identifikasi algoritma *Random Forest* tidak dapat mengungguli hasil identifikasi berbasis kombinasi indeks spektral dapat disebabkan oleh kesalahan interpretasi saat proses penentuan *Training Sample*. Sedangkan hasil identifikasi berdasarkan metode kombinasi indeks spektral (UI-NDVI-MNDWI) mendapat hasil terbaik karena proses perhitungan nilai pikselnya menggunakan indeks spektral yang telah disegmentasi menggunakan *Otsu Thresholding*, yang mana pembagian kelas nilai piksel sudah jelas antara obyek dan yang bukan obyek. Hal tersebut menambah akurasi dalam proses identifikasi lahan terbangun yang terbebas dari vegetasi dan badan air.

Berdasarkan data Bappeda Kota Surakarta Tahun 2021, didapatkan data luas lahan terbangun di Kota Surakarta sebesar 4161,42 Ha sedangkan non lahan terbangun adalah 465,14 Ha. Jika dihitung korelasi antara data tersebut dengan Luas lahan terbangun hasil identifikasi kombinasi indeks spektra (UI-NDVI-MNDWI) di tahun yang sama, maka didapat nilai R^2 sebesar 1. Artinya kedekatan antara nilai luas hasil identifikasi berbasis kombinasi indeks dengan luas dari data Bappeda sangat tinggi. Sehingga metode identifikasi obyek lahan terbangun berbasis kombinasi indeks spektral (UI-NDVI-MNDWI) menggunakan *Google Earth Engine* dapat digunakan secara kontinyu untuk memantau pola perubahan lahan di Kota Surakarta secara multi-temporal.



4. SIMPULAN

Hasil identifikasi lahan terbangun menggunakan metode klasifikasi terbimbing *Random Forest* dan identifikasi berbasis kombinasi indeks spektral (UI-NDVI-MNDWI) menunjukkan pola yang searah. Luas lahan terbangun tahun 2019 hingga tahun 2023 yang didapat dari ketiga metode tersebut selalu menunjukkan peningkatan nilai. Dari hasil uji akurasi, metode terbaik dari ketiga metode tersebut adalah identifikasi lahan terbangun berbasis kombinasi indeks spektral UI-NDVI-MNDWI dengan nilai rata-rata OA sebesar 98,047% dan Kappa 0,969. Sehingga metode tersebut dapat diandalkan untuk interpretasi obyek lahan terbangun dan dapat digunakan secara kontinyu untuk memantau pola perubahan lahan di Kota Surakarta.

DAFTAR PUSTAKA

- [1] Puspitasari, S., & Suharyadi. (2016). Kajian Kepadatan Bangunan Menggunakan Interpretasi Hibrida Citra Landsat-8 OLI di Kota Semarang Tahun 2015. *Jurnal Bumi Indonesia*, 5, 1-9.
- [2] BPS Pemerintah Kota Surabaya. (2016). Surakarta dalam angka 2022. Badan Pusat Statistik Kota Surakarta Tahun 2022.
- [3] Shelestov, Andrii, Mykola Lavreniuk, Nataliia Kussul, and Alexei Novikov. 2017. "Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping." 5(February): 1-10. <https://doi.org/10.3389/feart.2017.00017>
- [4] Mutanga, Onesimo, and Lalit Kumar. 2019. "Google Earth Engine Applications." *Remote Sensing* 11(591): 11-14. <https://doi.org/10.3390/rs11050591>
- [5] Teluguntla, P., & Thenkabail, P. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS J. Photogramm. Remote Sens.*, 325-340
- [6] Kurniawan, S., Nurhaidar, W. O., & Salihin, I. (2017). Optimalisasi Transformasi Spektral UI, NDBI, NDVI dan Kombinasi Transformasi Spektral UI-NDVI dan NDBI-NDVI Guna Mendeteksi Kepadatan Lahan Terbangun di Kota Magelang. *Jurnal Geografi Aplikasi Dan Teknologi*, 1 (1), 13-22.
- [7] Hidayati, I. N., Suharyadi, R., & Danoedoro, P. (2018b). Kombinasi Indeks Citra untuk Analisis Lahan Terbangun dan Vegetasi Perkotaan. *Majalah Geografi Indonesia*, 32 (1), 24. <https://doi.org/10.22146/mgi.31899>
- [8] Mustofa, Dian (2018). Perbandingan Metode Klasifikasi Berbasis Machine Learning Pada Google Earth Engine Untuk Pemetaan Perubahan Penutup Lahan (Studi Kasus: Daerah Aliran Sungai Opak-Oyo). Diunduh dari <http://etd.repository.ugm.ac.id/>.
- [9] Kamal, M., Farda, N. M., Jamaluddin, I., Parel, A., Wikantika, K., Prasetyo, L. B., & Irawan, B. (2020). A preliminary study on machine learning and google earth engine for mangrove mapping. *IOP Conference Series: Earth and Environmental Science* (Vol. 500, No. 1, p. 012038). IOP Publishing.