EnSTaR

Environmental Sensing Technologies and Research

Eramurra mapping
Algal mat - Remote Sensing

Technical report for:
Leichhardt
A7/436 Roberts Road
Subiaco, WA 6008

EnSTaR
ABN: 79 861 342 089
41 Hilltop Place, Kelmscott, WA, 6111, AUSTRALIA

Document Information

Report No.	RL004c
Title	Eramurra Mapping Algal Mat - Remote Sensing
Date	January 2024
Job Reference	LEI-PO-0029 PO from Leichhardt Salt Pty Ltd
Job Description	Process Sentinel satellite data to produce maps of environmental regions as defined by Leichhardt.

Version 1 submitted.	November 2022
Addendum produced. Project development envelopes changed slightly, requiring recalculation of some parameters.	July 2023
Final version. Changes in response to EPA feedback. Removal of project envelopes data after discussion with Leichhardt.	Jan 2024

DISCLAIMER

Information and data provided in this technical report are provided solely for the user's information and, while considered to be accurate at the time of publishing, are provided strictly as the best understanding of EnSTaR and without warranty of any kind. No warranty or guarantee, whether express or implied, is made with respect to the data reported or to the findings, observations and conclusions expressed in this report. EnSTaR, its agents, employees or contractors will not be liable for any damages, direct or indirect, or lost profits arising out of the use of information or data provided in this report or as associated digital media beyond their immediate implications.

1. Background

Leichhardt Salt Pty Ltd has interests in exploring the feasibility of satellite remote sensing data from Sentinel-2 A/B in mapping the algal mat in the northwest of Western Australia. The algal mat maps and data derived from this feasibility study are to support studies of the current environmental state, and change from recent times, with relevance to potential industrial development in the region.

Algal mats are characterised by the presence of productive microbial communities at or near the surface of marine sediments, ranging from subtidal to supratidal environments. Sometimes the microbial concentration is obvious as colouration at the surface of the sediment, but sometimes the microbial concentration is active with the sediment profile, therefore not particularly apparent as sediment discolouration.

A conventional field-based survey in mapping algal mats is accurate and useful in identifying algal mats in a small and target region, but it is ineffective and cost probative in a large area of observation. The aerial and satellite data provide an alternative to field-based surveillance in mapping algal mats for large spatial extents. The machine learning approaches, such as Support Vector Machine (SVM) and Random Forest (RF) classification in mapping land cover types have gained momentum in the last decade (Jiang et al. 2021). The rapid adoption of machine learning approaches is because of access to large processing/compute power readily available via various cloud computing providers such as Amazon Web Service and Google Earth Engine (Diniz et al. 2019) and machine learning approaches are efficient in mining and using spatial information (Dennis et al. 2012).

In this study, we used the SVM model for algal mats classification with the aim to:

1. Study the feasibility of satellite remote sensing data from Sentinel-2A/B in mapping algal mats in the study area shown in Figure 1.
2. Develop a methodology and consider limitations to identify algal mats from the satellite data and generate an algal mat map to inform the state of the algal mat in the study regions.

2. Study Sites

The region of interest, shown in Figure 1, extends from the southern extent of Exmouth Gulf to approximately 118 degrees east. A localized region representing the focus of the proposed development is indicated in Figure 1 by a magenta rectangle (Eramurra AOI). A larger project

area is indicated by a cyan boundary (Horseflats AOI). The project also has an interest in the broader regional area bounded by blue in Figure 1 (Regional AOI).

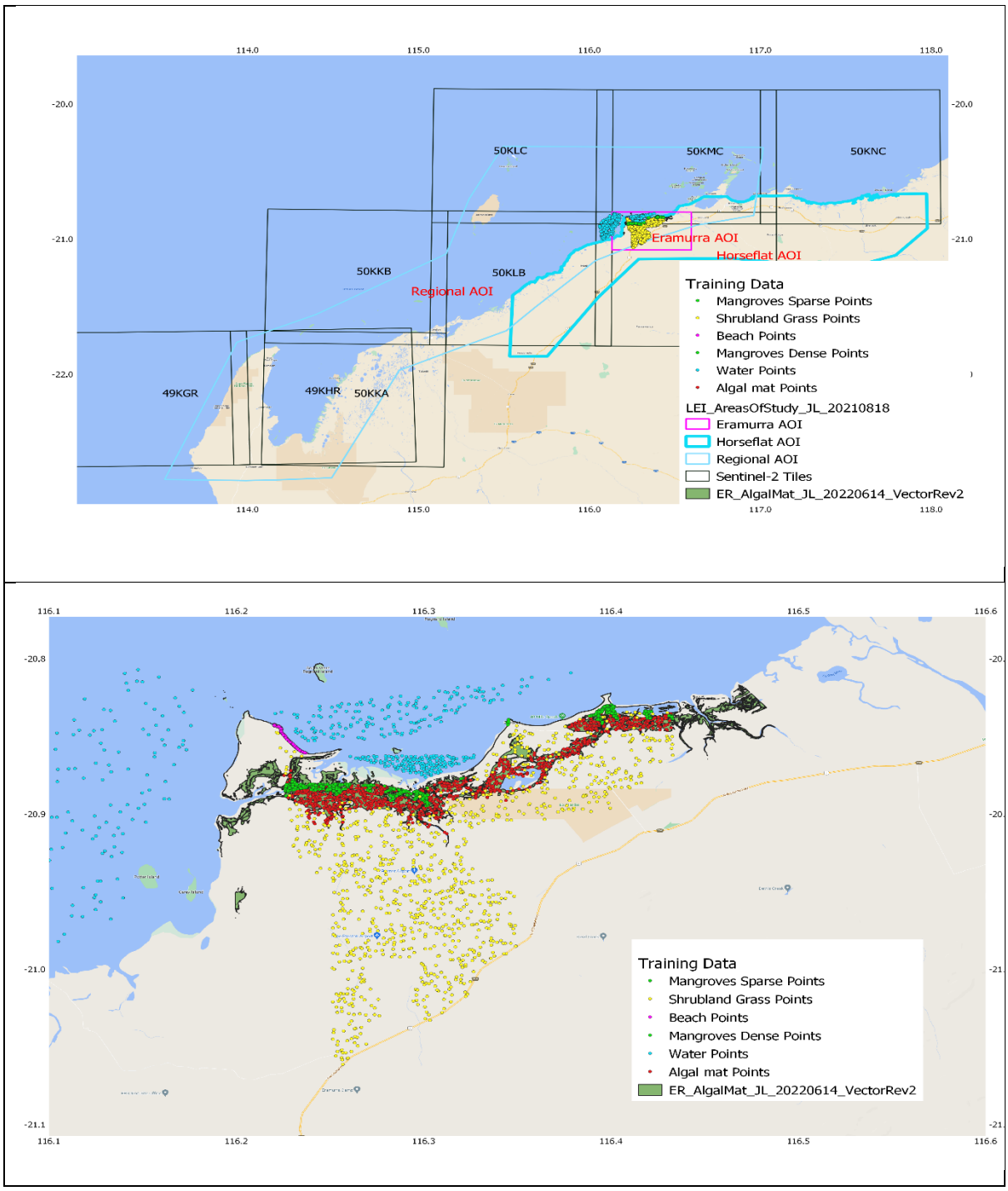

Figure 1. Top: Region of interest for remote sensing-based mapping comprises areas bounded by blue, cyan, and magenta. Bottom: The training data selection points provided by Leichhardt Salt Pty Ltd over the Eramurra AOI.

3. Data

The optical satellite data from Sentinel-2 A/B (See Section 3.1 for the details) and Synthetic Aperture Radar (SAR) data from Sentinel-1 A/B from the European Space Agency were used in this study. The SAR dataset used comprised of 2016 composite of normalized backscatter data in VV polarization (Bauer et al. 2021). The 1-second STRM Digital Elevation Model (DEM) from Geoscience Australia was also used. In addition, the ground coverage of the algal mat in the study area provided by Leichhardt Salt Pty Ltd was used in selecting the training points for the SVM model and validating the results.

3.1 *Satellite Remote Sensing Data and Indices*

The Sentinel-2A and 2B each provide 10-day repeat views of the globe, interleaved to provide views of the earth at intervals of 5 days. Sentinel 2A began operation in June 2015 and Sentinel 2B in March 2017. Spatial resolution varies from 10 m to 60 m across 13 visible and infrared spectral bands (see Table 1). For this work, we used the Sentinel-2 data from 1st January 2018 to 31st of December 2021 for Algal mat mapping. The Normalized Bottom of Atmosphere Reflectance and Terrain corrected (NBART) product from Sentinel-2 A/B and cloud mask generated using python implementation of the `fmask` algorithm by Zu and Woodcock (2012) and Zhu et al. (2015). The NBART products are preferred over the top of the atmosphere reflectance because NBART products have the atmospheric and terrain effects corrected to account for atmospheric, sun and satellite angles. The NBART product allows for accurate comparison of imagery at different product locations and seasons because inconsistencies can arise between the satellite images at different time periods because of variations in atmospheric conditions, sun and satellite angles and terrain slope and aspects (Li et al. 2010).

Table 1. Sentinel-2 Multi-spectral Instrument (MSI) band information.

	wavelength	resolution (m)
Band 1 – Coastal aerosol	443	60
Band 2 – Blue	492	10
Band 3 – Green	560	10
Band 4 – Red	665	10
Band 5 – Vegetation red edge	704	20
Band 6 – Vegetation red edge	740	20
Band 7 – Vegetation red edge	783	20
Band 8 – NIR	833	10
Band 8A – Narrow NIR	865	20
Band 9 – Water vapour	945	60
Band 10 – SWIR – Cirrus	1373	60
Band 11 – SWIR	1614	20
Band 12 – SWIR	2202	20

3.2 *Annual Median Sentinel-2 Derived Indices*

The dominant land systems in the study region in the coastal areas comprise sandy beaches, mudflats, mangroves, and sparse vegetation. Towards the inland region, the land systems comprise sparse vegetation shrublands and red earth. The Red Soil Index (RSI), NDVI, Enhanced Vegetation Index (EVI) and NDWI derived from Sentinel-2 were used as input to the SVM model.

Table 2. Sentinel-2 indices that were used were used as an input to the SVM Model.

Enhanced Vegetation Index (EVI)	$EVI = 2.5 * \frac{Band8 - Band4}{(Band8 + 6.0 * Band4 - 7.5 * Band2) + 1}$	Huete et al. (2002)
Normalised Difference Vegetation Index (NDVI)	$NDVI = \frac{Band4 - Band8}{Band4 + Band8}$	Rouse et al. (1974)
Normalised Difference Water Index (NDWI)	$NDWI = \frac{Band3 - Band8}{Band3 + Band8}$	Gao (1996)
Red Soil Index (RSI)	$RSI = \frac{Band4}{Band2}$	

The indices listed in Table 2 were computed for each pass individually after masking the clouds. Because of the low temporal resolution (~5 days) of Sentinel-2 A/B data, pixels identified as a cloud meant the data was missing. To address the missing data due to clouds

and to normalize a seasonal anomaly, a yearly median composite of indices was generated. For a specific site, the median of 3 x 3 pixels was extracted from the index data and the median of 3 x 3 pixels was selected to account for the fact that the pixels are affected by the adjacent pixels.

3.3 *In situ Data*

The pre-identified algal mat features vector data provided by Leichhardt Salt Pty Ltd are used in this study to select the training points of the algal mat for the SVM model and validate the classification results. In total, 2554 polygons and 120 points were provided with different algal mat classification types. Table 3 lists the classification types in the polygons and Table 4 for the point datasets. All the sites were located within the area identified as an Eramurra AOI (see Figure 1-bottom).

Table 3. The algal mat classification type and the number of features identified in each class for polygon datasets.

Number	Algal mat Class / Description	No of polygons
1	Low near	539
2	Low far	1393
3	Limited activity	466
4	Active mat	156

Table 4. The algal mat classification type and the number of features identified in each class for point datasets

Number	Algal mat Class / Description	No of Points
1	Algal	94
2	Algal and Acid Sulphate	23
3	Grab	3

3.4 *Training Datasets for the SVM Model*

The training datasets for the SVM model were collected after overlaying the algal-mat polygons produced by Phoenix Environmental Sciences and provided by Leichhardt Salt Pty Ltd. A total of 1000 algal mat points (See Figure 1 (bottom) red points) were extracted from the regions inside the algal mat extents and visual interpretation of high-resolution Google Earth imagery. Further, 3221 points were selected from other regions that were representative

of other surface types listed in Table 5 after overlaying with polygons of different land types. Figure 1 (bottom) shows the locations of the training data points.

Table 5. The number of points from different surface types used in selecting the training datasets.

Number	Land Types	No. of Training Points
1	Algal mat Points	1000
2	Mangroves sparse Points	500
3	Mangroves dense Points	500
4	Shrubland Grass Points	1000
5	Water Points	480
6	Beach Points	741

4. Methodology and Results:

4.1 *SVM Model*

Further, to help provide some context for inland land types, we included the DEM and Sentinel-1 backscatter data in VV polarization described in Section 3.

From the 2020 and 2021 annual median indices datasets, we extracted a median of 3 x 3-pixel values across all the training sites. Only yearly median indices from 2020 and 2021 were used in selecting the training datasets because the polygons overlaid to extract the points were from 2020 and 2021.

In total, 2,000 points were datasets from the algal mat and 6,442 points from other surface types were extracted as input for training the SVM model.

In training the SVM model, we used 70% of the datasets, while 30% were retained for validation of the model. The validation results of the SVM model had an Accuracy of 94% kappa coefficient of 90.02% with a precision of 95% and a recall rate of 94% with F1-score of 94.52%.

4.2 SVM Model *validation using in situ algal points*

A total of 120 *in situ* algal mat points were available in the point validation dataset. There were only 67 validation points left after filtering and selecting only points that were confirmed to have algal materials via the remark in the validation datasets remark section. The point

sampling validation for SVM model generated algal mat map was done by identifying if the validation points were identified as algal mat or not.

Table 6. Validation results for true positive identification of algal mat.

Year	No of True Positive Detection	Accuracy (%)	F1-Score (%)
2018	64	95.52	97.70
2019	63	94.03	96.92
2020	62	92.54	96.12
2021	62	92.54	96.12

4.3 SVM Model v*alidation using in situ vectorial features*

The algal mats classified by the SVM for the Eramurra region are presented from Figure 2 to Figure 5 for the years 2018 – 2021 respectively. We observe that the vectorial layer mapped as an algal mat is mostly classified as an algal mat by the SVM model. Note, the four colours representing the vector polygons produced by Phoenix Environmental Sciences are overlayed on the SVM grey region, but are also transparent, thus appear subdued in brightness. However, there are some areas, most notably bright green, where the SVM model has not identified the pixels as an algal mat.

Table 7 presents the percentage of pixels the SVM model correctly identified from the different features in the *in situ* vector datasets. The percentage of pixels correctly identified within the vector features ranged from 77.23% to 97.26%. The pixels in the `Low far` feature type were poorly classified by the SVM model when compared with other algal mat feature types. The `Low far` features were located primarily around the vicinity of mangroves.

Table 7. The number of pixels identified correctly (expressed in percentage) as algal mat within each algal mat classification in in situ vector datasets.

Year	Accuracy (%)			
	Low near	Low far	Limited activity	Active mat
2018	77.23	95.24	94.36	97.13
2019	77.71	96.14	94.68	97.26
2020	74.61	91.43	93.94	96.52
2021	73.49	90.34	94.18	96.87

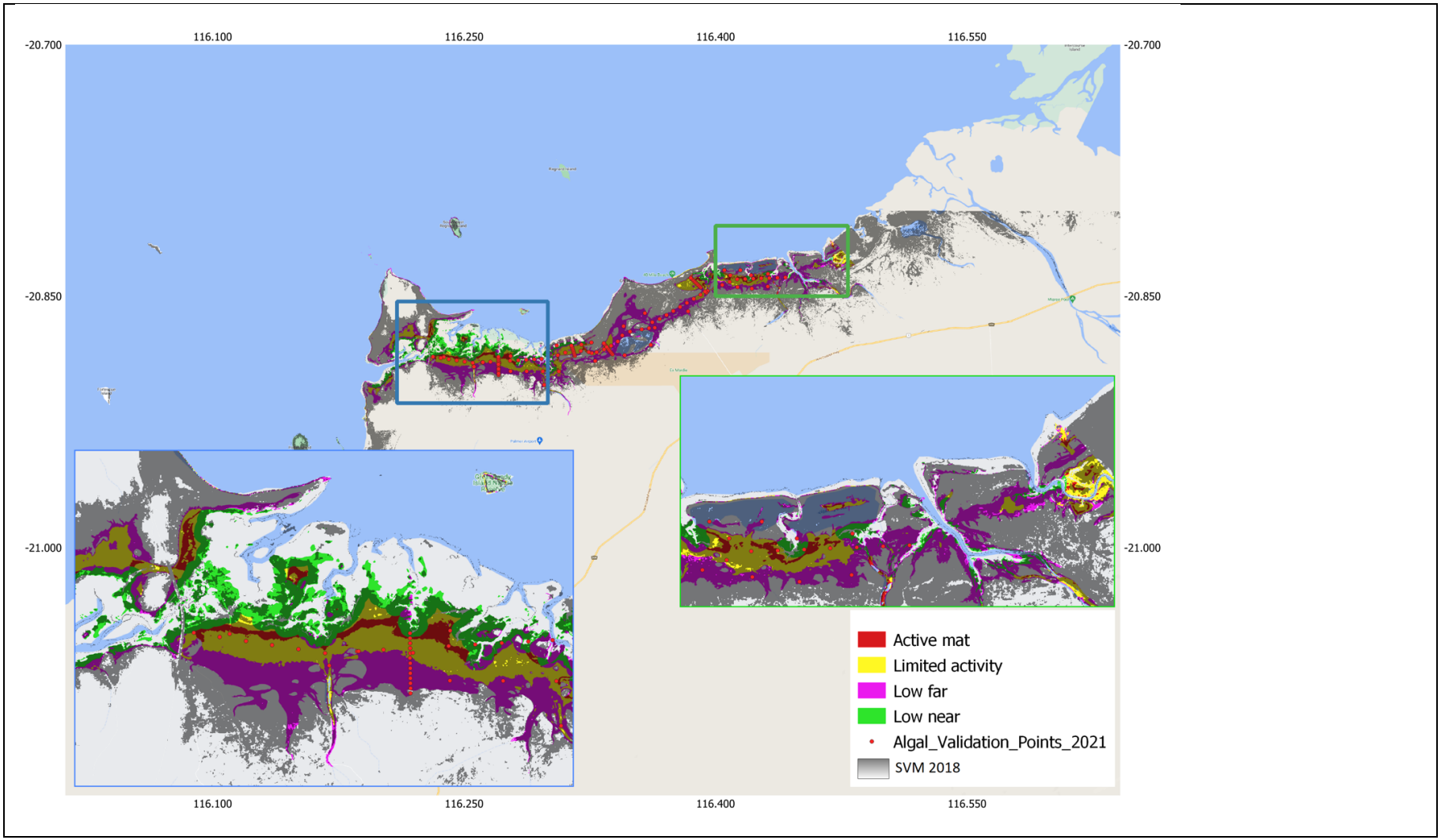

Figure 2. Algal mat classification result generated by the SVM Model for the year 2018 for the Eramurra Area of Interest (grey regions). The coloured regions represent the classification polygons produced by Phoenix Environmental Sciences.

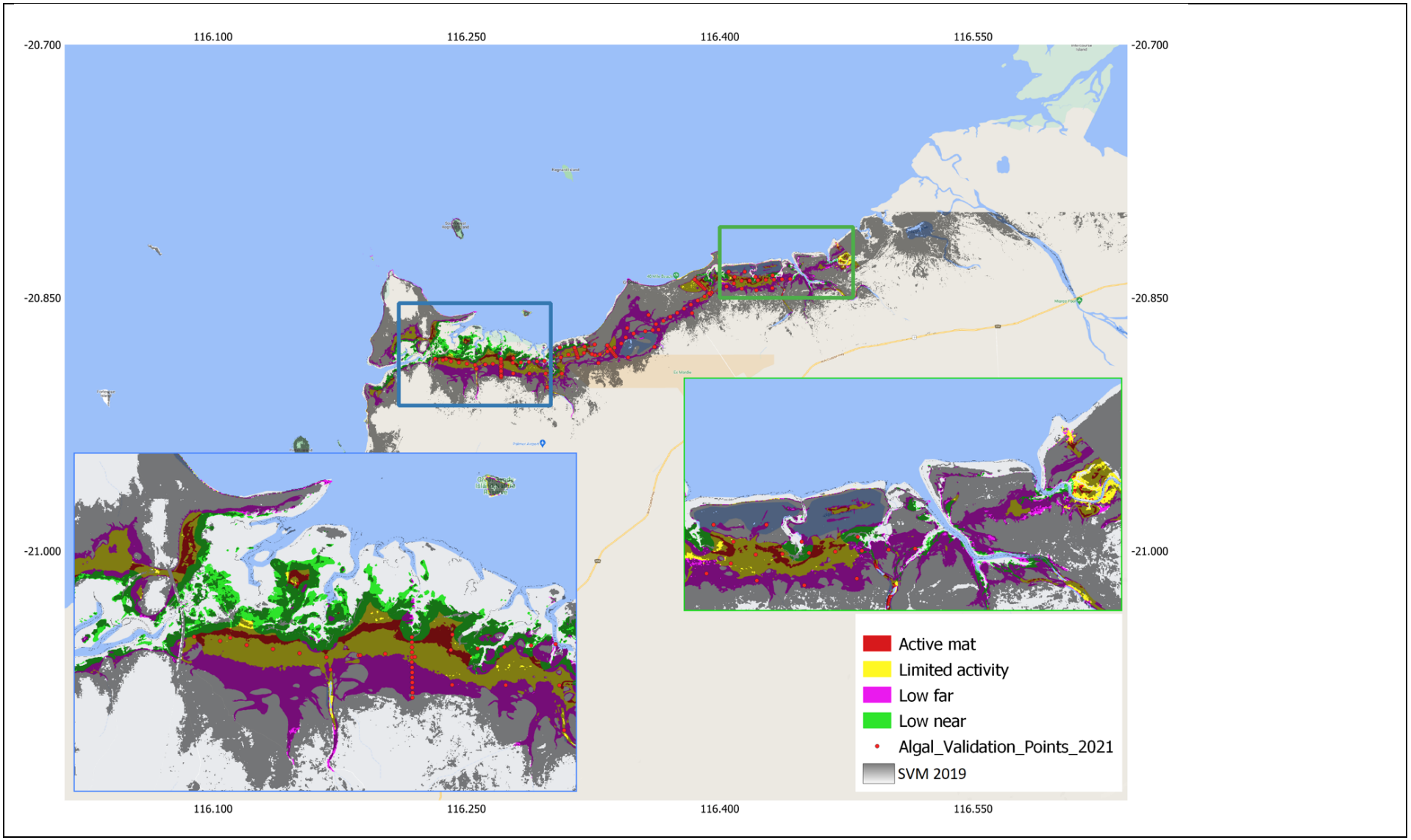

Figure 3. Algal mat classification result generated by the SVM Model for the year 2019 for the Eramurra Area of Interest (grey regions). The coloured regions represent the classification polygons produced by Phoenix Environmental Sciences.

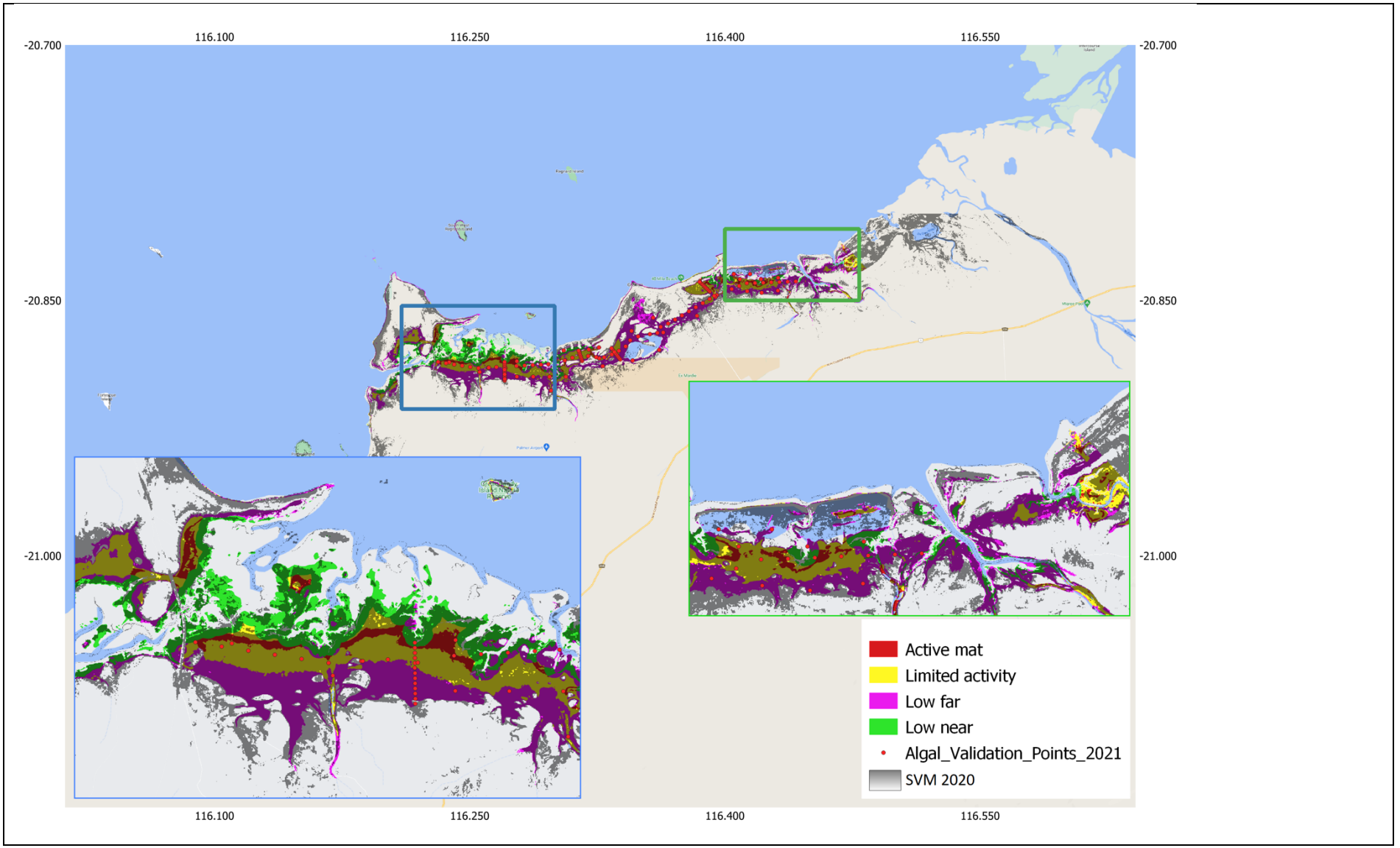

Figure 4. Algal mat classification result generated by the SVM Model for the year 2020 for the Eramurra Area of Interest (grey regions). The coloured regions represent the classification polygons produced by Phoenix Environmental Sciences.

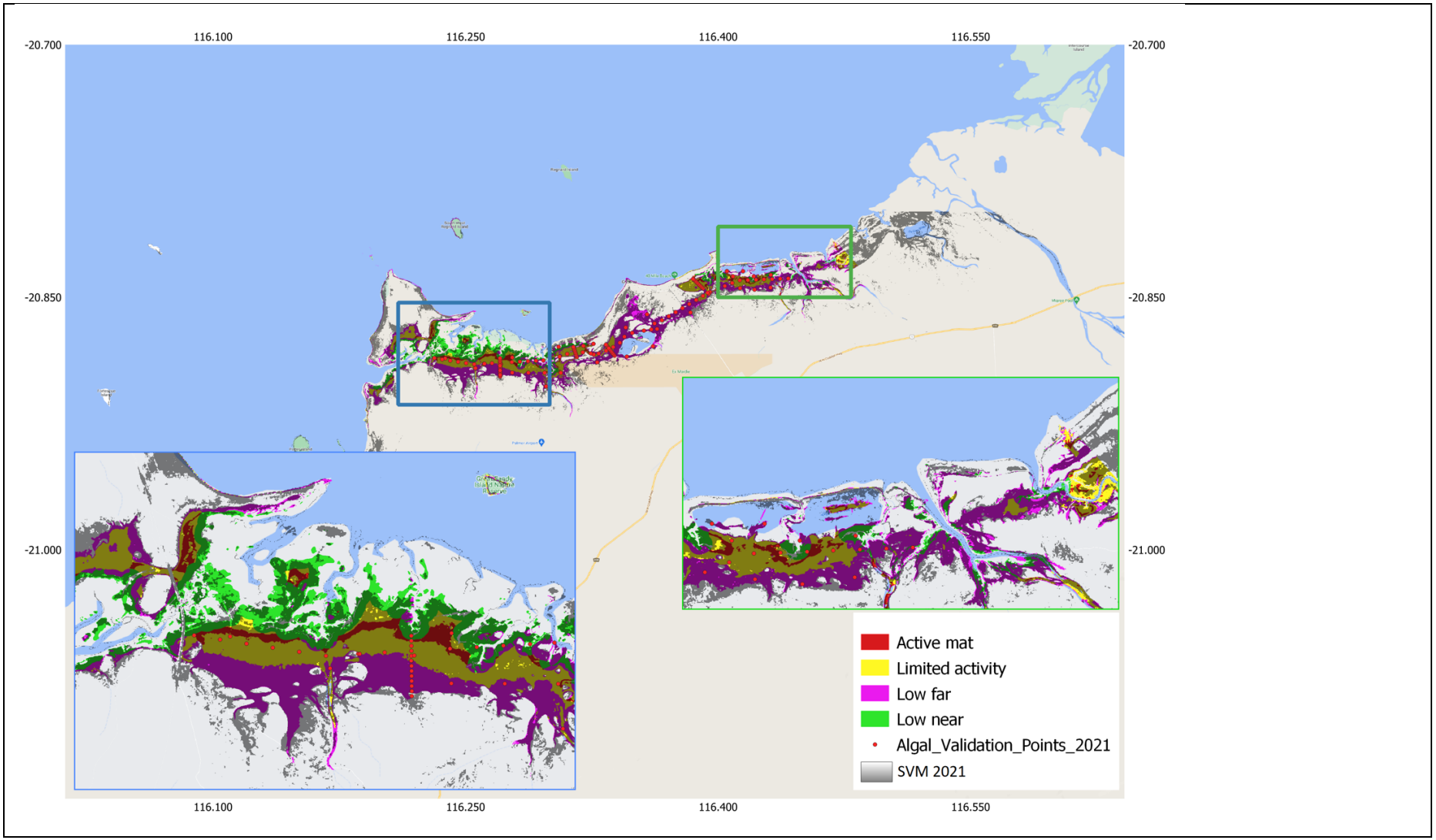

Figure 5. Algal mat classification result generated by the SVM Model for the year 2021 for the Eramurra Area of Interest (grey regions). The coloured regions represent the classification polygons produced by Phoenix Environmental Sciences.

4.4 *SVM classified algal mats*

The total area classified as algal mat by the SVM model is presented in Table 8 for the region bounded by all the sentinel-2 tiles in Figure 1.

Table 8. Total area classified as algal mat by the SVM process for years 2018 to 2021.

Year	**Total Area (km^2)**
2018	3101
2019	3187
2020	3126
2021	2705

4.5 SVM classified productive algal mats – full coastal extent

The spatial extent of validation/training data is confined to the regions depicted in Figure 1. Development of a robust algal mat classification scheme for regions far-removed from the training region must be approached with caution. Nonetheless, we have explored a simple approach to identifying two classes of algal mat based on NDVI, a remote sensing product related to productivity. For this work, we selected NDVI thresholds of Max NDVI >= 0.4 and STD NDVI >= 0.04 to classify algal mat pixels as productive. Only pixels that meet both thresholds are deemed “productive algal mat”. Pixels that fail one or both thresholds are considered low productivity. Figures 6 to 10 show the SVM derived algal mat regions for years 2018 to 2021. Algal mat regions are displayed as black. Pixels classed as productive are highlighted in orange.

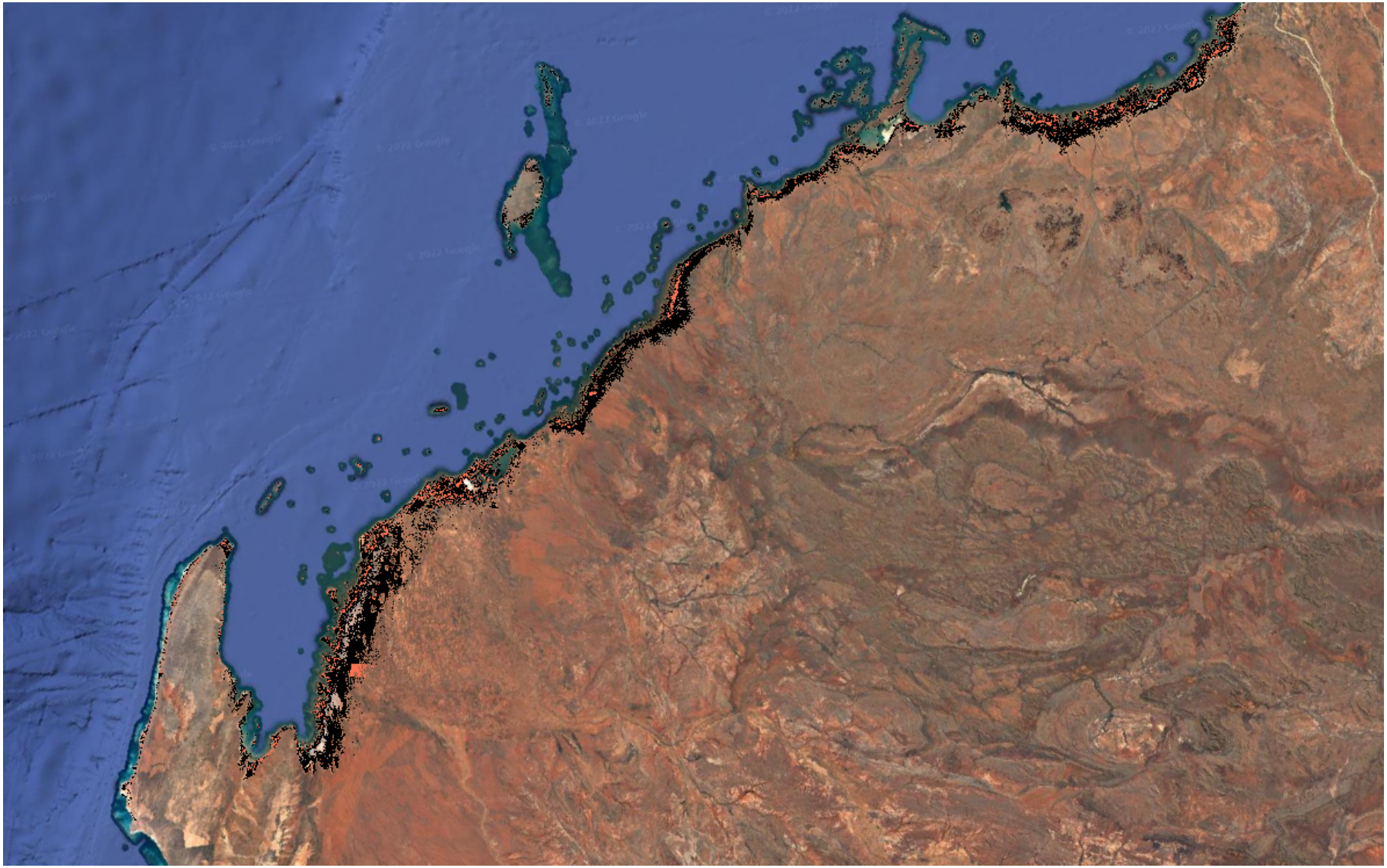

Figure 6. 2018 SVM derived algal mat (black) and productive algal mat (orange) based on MAX NDVI >=0.4 and STD NDVI >= 0.04.

Figure 7. 2019 SVM derived algal mat (black) and productive algal mat (orange) based on MAX NDVI >=0.4 and STD NDVI >= 0.04.

Figure 8. 2020 SVM derived algal mat (black) and productive algal mat (orange) based on MAX NDVI >=0.4 and STD NDVI >= 0.04.

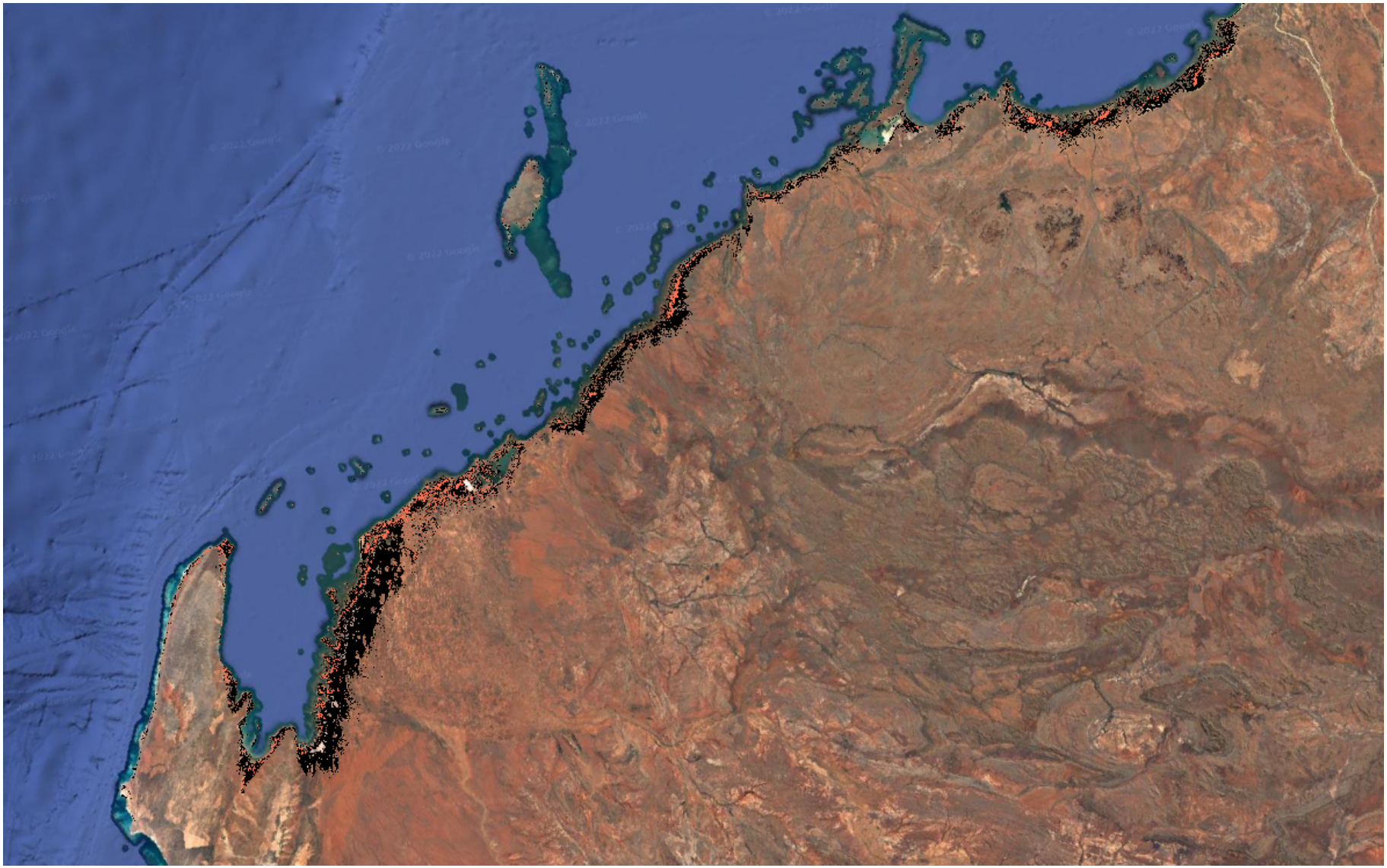

Figure 9. 2021 SVM derived algal mat (black) and productive algal mat (orange) based on MAX NDVI >=0.4 and STD NDVI >= 0.04.

4.5 *Productivity-based algal mat classification – Eramurra Area of Interest*

For algal mat classification we grouped *in situ* algal mat vectorial features listed in Table 3 into three groups: "Active mat," "Limited activity," and "Low activity." The "Low activity" group is a combination of "Low near" and "Low far" vectorial features.

Since NDVI values are directly correlated to algal mat productivity, the thresholds derived from the maximum NDVI values were used to assign algal mats generated by the SVM model into one of the three classes described above.

The thresholds were derived from the maximum NDVI values computed for the year 2021 (as 2021 is the closest year to 2022 when *in situ* algal mat features were created). All permutations of different statistical values presented in Table 9 were tested as possible threshold values and validated using 2063 random validation points selected from inside polygons from the three classes described above. The top five results of maximum NDVI thresholds for multi-class algal mat validations are presented in Table 10. From the results presented in Table 10, we defined the maximum NDVI threshold as:

1. "Active mat" >= 0.566
2. 0.255 <= "Limited activity" > 0.566
3. 0.105 <= "Low activity" > 0.255

To classify algal mat pixels into "Active mat", "Limited activity", and "Low activity" classes, we applied a condition where pixels that were previously classified as an algal mat by the SVM model must have yearly maximum NDVI values within the threshold values defined above.

Table 9. Statistical values derived from maximum NDVI data for the year 2021 for different algal mat classes.

Statistics	Max	Min	Mean	Percentiles									
Class				2.5	5	10	25	35	45	50	60	75	90
Active mat	0.70	0.029	0.43	0.15	0.16	0.18	0.29	0.39	0.45	0.48	0.52	0.57	0.60
Limited activity	0.83	0.023	0.34	0.12	0.14	0.16	0.20	0.26	0.31	0.34	0.39	0.45	0.53
Low activity	0.74	0.045	0.20	0.11	0.11	0.12	0.14	0.15	0.16	0.17	0.18	0.23	0.32

Table 10. NDVI thresholds with accuracy statistics for algal mat (which is a combination of "Active mat", "Limited activity", and "Low activity" class) for the five best results (Bold values were selected as the thresholds to differentiate different algal mat classes).

	Low activity NDVI Threshold	Limited activity NDVI Threshold	Active mat NDVI threshold	Accuracy	Recall	F1-Scores
1	**0.105**	**0.255**	**0.566**	0.891	0.887	0.942

2	0.105	0.255	0.478	0.887	0.887	0.940
3	0.105	0.255	0.518	0.887	0.891	0.940
4	0.105	0.255	0.598	0.887	0.887	0.940
5	0.105	0.400	0.478	0.887	0.887	0.940

The multi-class algal mat classifications for the Eramurra region are presented in Figures 10 to 13 for the years 2018 to 2021 respectively.

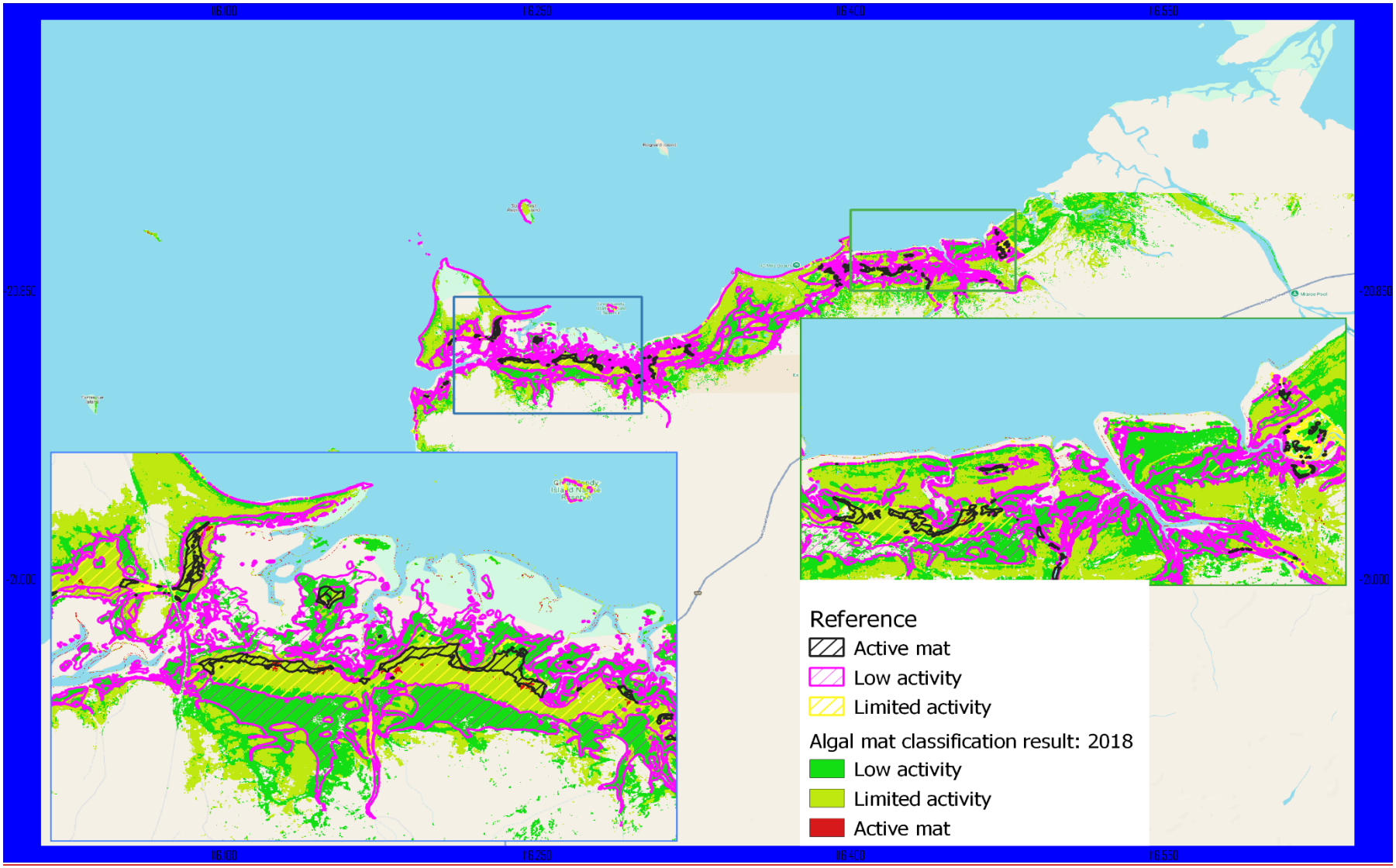

Figure 10. Multi-class algal mat classification result generated by using the SVM model data with maximum NDVI thresholds for the year 2018 for the Eramurra Area of Interest. Solid colours show the regions defined by the SVM Model. Hatched regions show the regions defined by the reference map (Reference is 2022 validation vectorial layer supplied by Leichhardt).

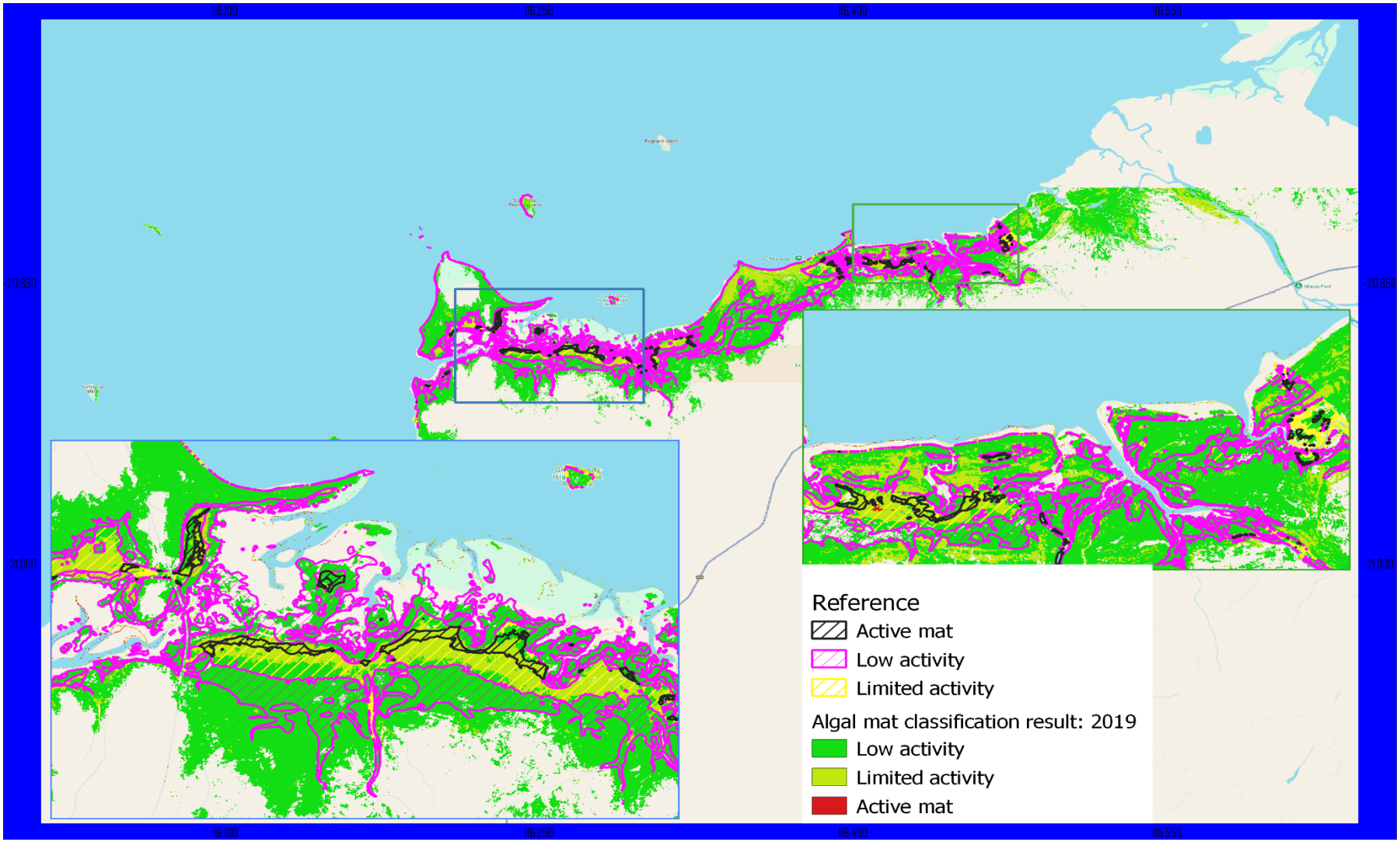

Figure 11. Multi-class algal mat classification result generated by using the SVM model data with maximum NDVI thresholds for the year 2019 for the Eramurra Area of Interest. Solid colours show the regions defined by the SVM Model. Hatched regions show the regions defined by the reference map (Reference is 2022 validation vectorial layer supplied by Leichhardt).

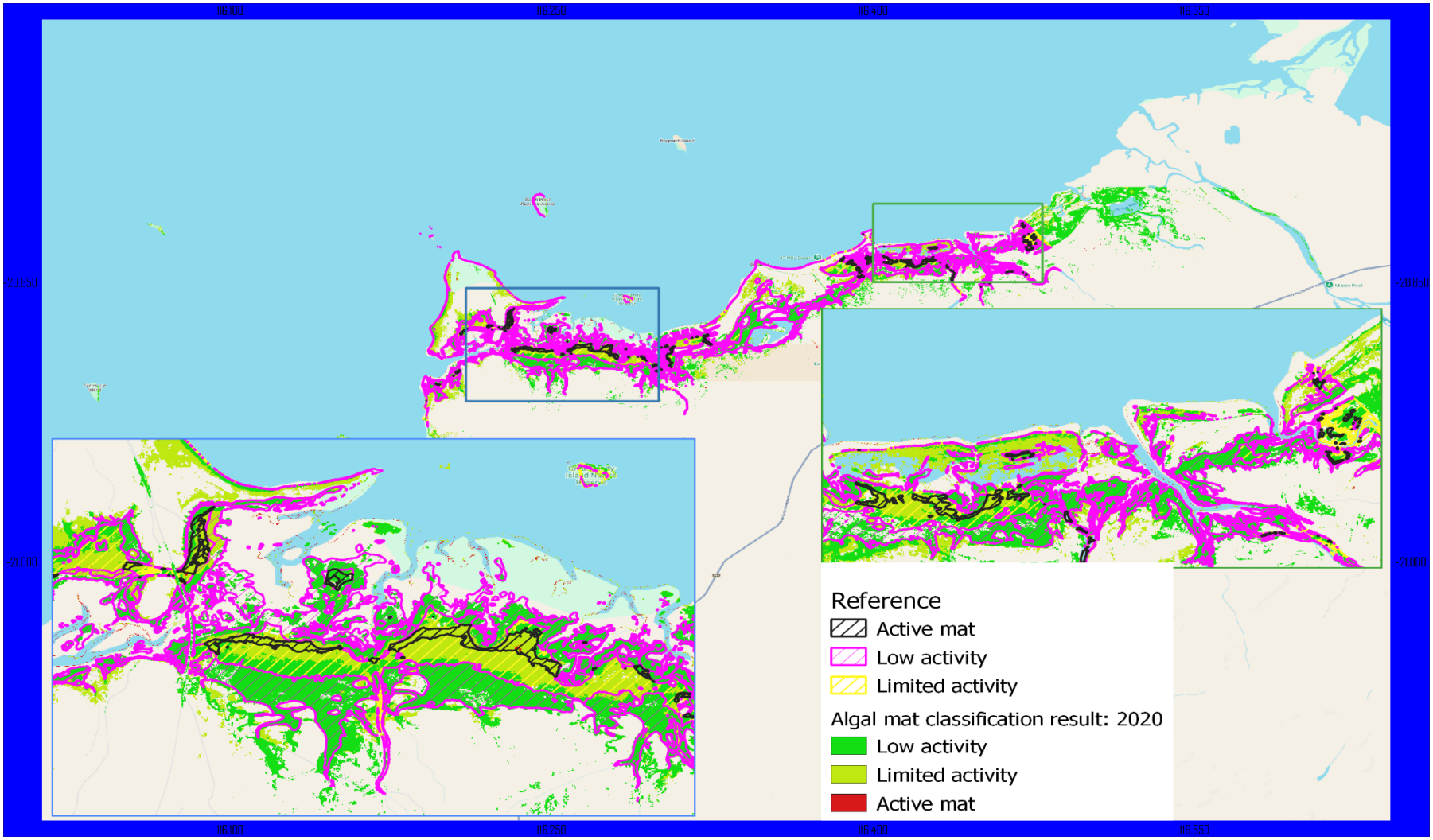

Figure 12. Multi-class algal mat classification result generated by using the SVM model data with maximum NDVI thresholds for the year 2020 for the Eramurra Area of Interest. Solid colours show the regions defined by the SVM Model. Hatched regions show the regions defined by the reference map (Reference is 2022 validation vectorial layer supplied by Leichhardt).

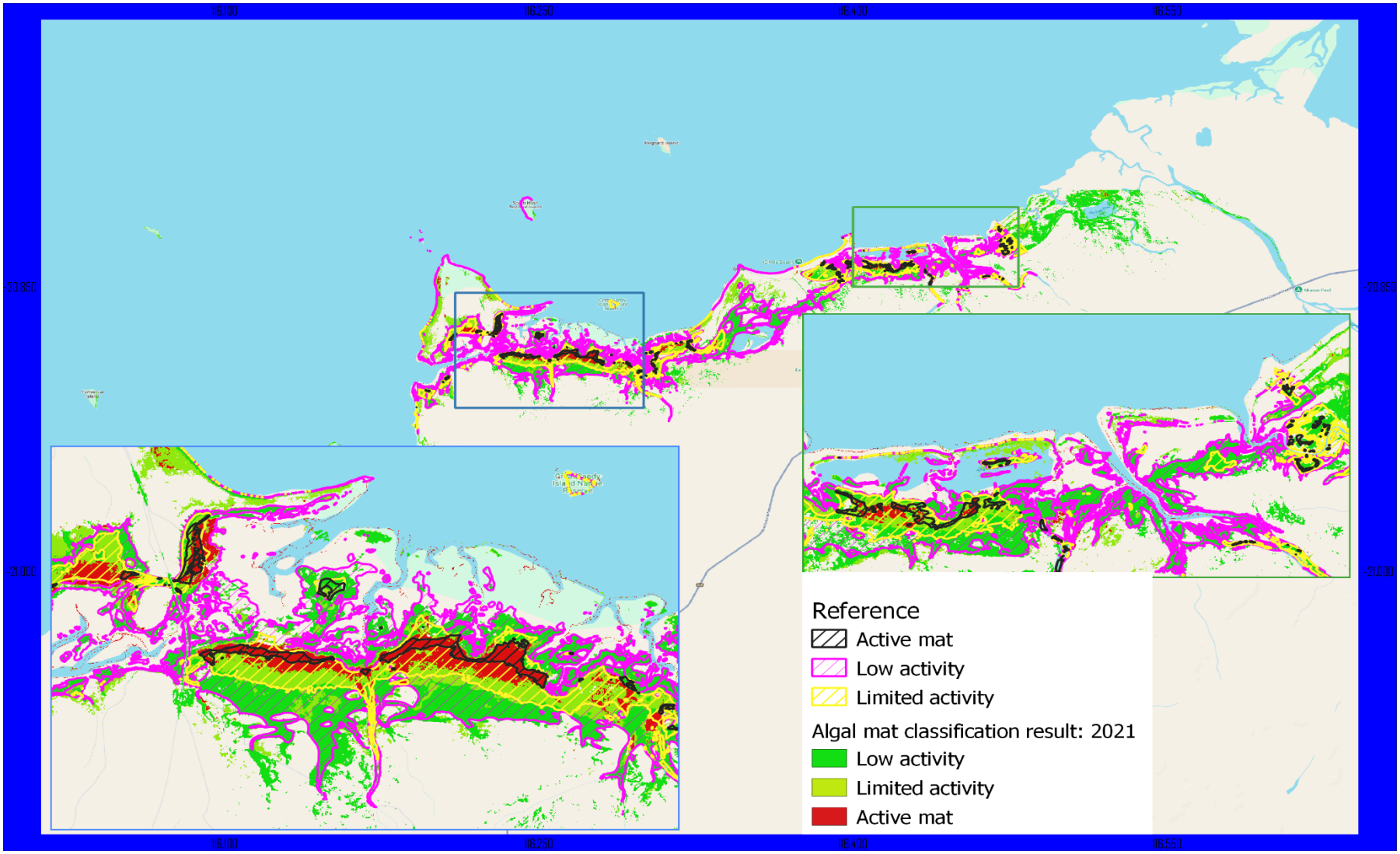

Figure 13. Multi-class algal mat classification result generated by using the SVM model data with maximum NDVI thresholds for the year 2021 for the Eramurra Area of Interest. Solid colours show the regions defined by the SVM Model. Hatched regions show the regions defined by the reference map (Reference is 2022 validation vectorial layer supplied by Leichhardt).

An indication of annual variability in the regional extent of algal mat, and the variability in the regional extent based on the SVM model, is presented in Table 11 which shows the total area of SVM-derived algal mat for the Eramurra AOI (see Figure 1).

Table 11. The total area of multi-class algal mat is based on the yearly NDVI maximum for the years 2018 to 2021. Note: total area is for the region enclosed in the Eramurra AOI in Figure 1.

Class/Year	Total Area **(km^2)**		
	Low activity region	Limited activity Region	Active mat region
2021	61.65	25.92	1.30
2020	76.62	31.91	0.29
2019	155.69	24.00	0.18
2018	91.00	69.52	0.41

6. Discussion

The mapping of algal mat using the SVM model was carried out over the north western region of Western Australia for the years 2018 to 2021 using a combination of optical remote sensing data from Sentinel-2 A/B, DEM, and Synthetic Aperture Radar data from Sentinel-1 A/B. Validation of the results was carried out based on data collected within the Eramurra Area of Interest (AOI). The validation analysis shows that the SVM model performed well in mapping the general extent of algal mat as well as providing an indication of the extent of productive algal mats.

The SVM model is "tuned" to the training data representing algal mat in the Eramurra AOI thus confidence in the map products further afield, such as towards the Exmouth Gulf, may be considered somewhat less than within the Eramurra AOI. The extensive maps presented in Figures 6 to 9 show a representative distinction between high and low productivity. Extensive and widespread field-based validation would be required to provide an indication of accuracy of map products outside the Eramurra AOI. It is interesting to note that, based on visual inspection of Figures 10 to 13, the SVM model appears to perform better in the southern extent of the Phoenix Environmental Sciences regions compared to the northern extent. There is a paucity of validation points in the northern extent.

We showed a method based on NDVI annual statistics and maximum values, to provide an indication of algal mat productivity. Comparison of the mapped SVM productive algal mat extent and the *in situ*-based maps of active mat are very encouraging. Users of the SVM method of algal mat mapping could consider more complex approaches to integrating NDVI values spatially and temporally to infer relative annual algal mat total regional productivity.

References

1. Bauer-Marschallinger, B., Cao, S., Navacchi, C. *et al.* The normalised Sentinel-1 Global Backscatter Model, mapping Earth's land surface with C-band microwaves. *Sci Data* **8,** 277 (2021). https://doi.org/10.1038/s41597-021-01059-7
2. Dennis, C.D.; Steven, E.F.; Monique, G.D. (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ*., 118, 259–272.
3. Diniz, C, Cortinhas, L, Nerino, G, Rodrigues, J, Sadeck, L, Adami, M & Souza-Filho, PWM (2019), 'Brazilian Mangrove Status: Three Decades of Satellite Data Analysis', *Remote Sensing*, vol. 11, pp.808
4. Gao, Bo-Cai (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*. **58** (3): 257–266
5. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.*, *83*, 195–213
6. Jiang, Y.; Zhang, L.; Yan, M.; Qi, J.; Fu, T.; Fan, S.; Chen, B. (2021) High-Resolution Mangrove Forests Classification with Machine Learning Using Worldview and UAV Hyperspectral Data. *Remote Sens*., 13, 1529
7. Li, F., Jupp, D. L. B., Reddy, S., Lymburner, L., Mueller, N., Tan, P., & Islam, A. (2010). An evaluation of the use of atmospheric and brdf correction to standardize landsat data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(3), 257–270. https://doi.org/10.1109/JSTARS.2010.2042281
8. Rouse, J.W, Haas, R.H., Scheel, J.A., and Deering, D.W. (1974) 'Monitoring Vegetation Systems in the Great Plains with ERTS.' *Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium*, vol. 1, p. 48-62.
9. Zhu, Z. and Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery Remote Sensing of Environment 118 (2012) 83-94.
10. Zhu, Z., Wang, S. and Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images *Remote Sensing of Environment* 159 (2015) 269-277