

Lithological mapping enhancement by integrating Sentinel 2 and gamma-ray data utilizing support vector machine: A case study from Egypt

Ali Shebl^{a,b,*}, Mahmoud Abdellatif^{c,d}, Musa Hissen^a, Mahmoud Ibrahim Abdelaziz^{c,e},
Árpád Csámer^{a,*}

^a Department of Mineralogy and Geology, University of Debrecen, 4032 Debrecen, Hungary

^b Department of Geology, Tanta University, 31527 Tanta, Egypt

^c Department of Geophysics, University of Miskolc, 3515 Miskolc, Hungary

^d Department of Geology, South Valley University, 83523 Qena, Egypt

^e Department of Geology, Mansoura University, 35516 Mansoura, Egypt

ARTICLE INFO

Keywords:

Lithological mapping
Sentinel 2
Spectrometric gamma-ray
Data fusion
Support vector machine

ABSTRACT

Hybrid data fusion mostly gives a better diagnosis to lithological units compared to single-source mapping techniques. Rock unit discrimination depends mainly on variations in the concentrations of chemical elements. Remote sensing datasets reflect these variations as different spectral reflectances, while gamma-ray spectrometric measurements enable recording the varied concentrations of K, Th, and U in these rock units. Accordingly, in this study, we use Support-Vector Machine (SVM) learning algorithm to classify combined high spectral resolution Sentinel 2 data with K, Th, and U content of the rocks to better differentiate a lithologically complex area in Egypt. SVM classifier has been trained and tested on a reference map (built from FCCs, principal and independent component analysis of remote sensing images, as well as previous geological maps) to allocate 13 lithological targets. K, Th, U, and total count maps are interpolated using the inverse distance weighted (IDW) method, cubically resampled, and fused with Sentinel 2 data. We concluded that incorporating any single chemical concentration in the allocation gives better results than using remote sensing data solely and raised the Overall Accuracy by 4.14%, 5.11%, and 6.83% by adding U, K, and Th, respectively. Moreover, blending the total count band (K + Th + U) with Sentinel 2 data outstandingly boosts the classification accuracy by 7.77 %. We performed field reconnaissance to verify the classification results. The study demonstrates the effectiveness of integrating Sentinel 2 data with airborne geophysical spectrometric data, and the proposed approach may prove a more precise and sophisticated lithological map.

1. Introduction

Orogenic belts around the world mostly constitute a major source of mineral deposits (Deng et al., 2014) including rare earth elements (REEs) however, a comprehensive geochemical survey is always challenging due to their vast area and rugged topography (Cheng et al., 2021). The lack of geochemical data hinders more efficient lithological mapping, which is the solid base for mineral exploration. Airborne gamma-ray spectrometric data could provide a reasonable mapping over extensive terranes (Harris and Grunsky, 2015). These spectrometric data could be well-implemented in lithological identification and hydro-thermal alteration confirmation (Shebl et al., 2021) especially with the

availability and proven advancements of Machine learning algorithms (MLAs) and remote sensing datasets in geosciences (Harris and Grunsky, 2015; Pal and Mather, 2005).

Gamma-ray spectrometry provides a method for detecting the geochemical variations of potassium, thorium, and uranium in material near the earth's surface (Dickson and Scott, 1997). Various lithologies show different discrete radioelement concentrations. Consequently, the measured concentrations from the gamma-ray spectrometric data may be utilized to identify zones of the consistent lithology and contacts between contrasting lithologies (Anderson and Nash, 2018; Charbonneau et al., 1997), through (i) ternary map (Elkhateeb and Abdellatif, 2018; Patra and Veldi, 2016), (ii) statistical methods including cluster

* Corresponding authors at: Department of Mineralogy and Geology, University of Debrecen, 4032 Debrecen, Hungary. Department of Geology, Tanta University, 31527 Tanta, Egypt (A. Shebl).

E-mail address: ali.shebl@tanta.science.edu.eg (A. Shebl).

<https://doi.org/10.1016/j.jag.2021.102619>

Received 12 August 2021; Received in revised form 2 November 2021; Accepted 9 November 2021

Available online 12 November 2021

0303-2434/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

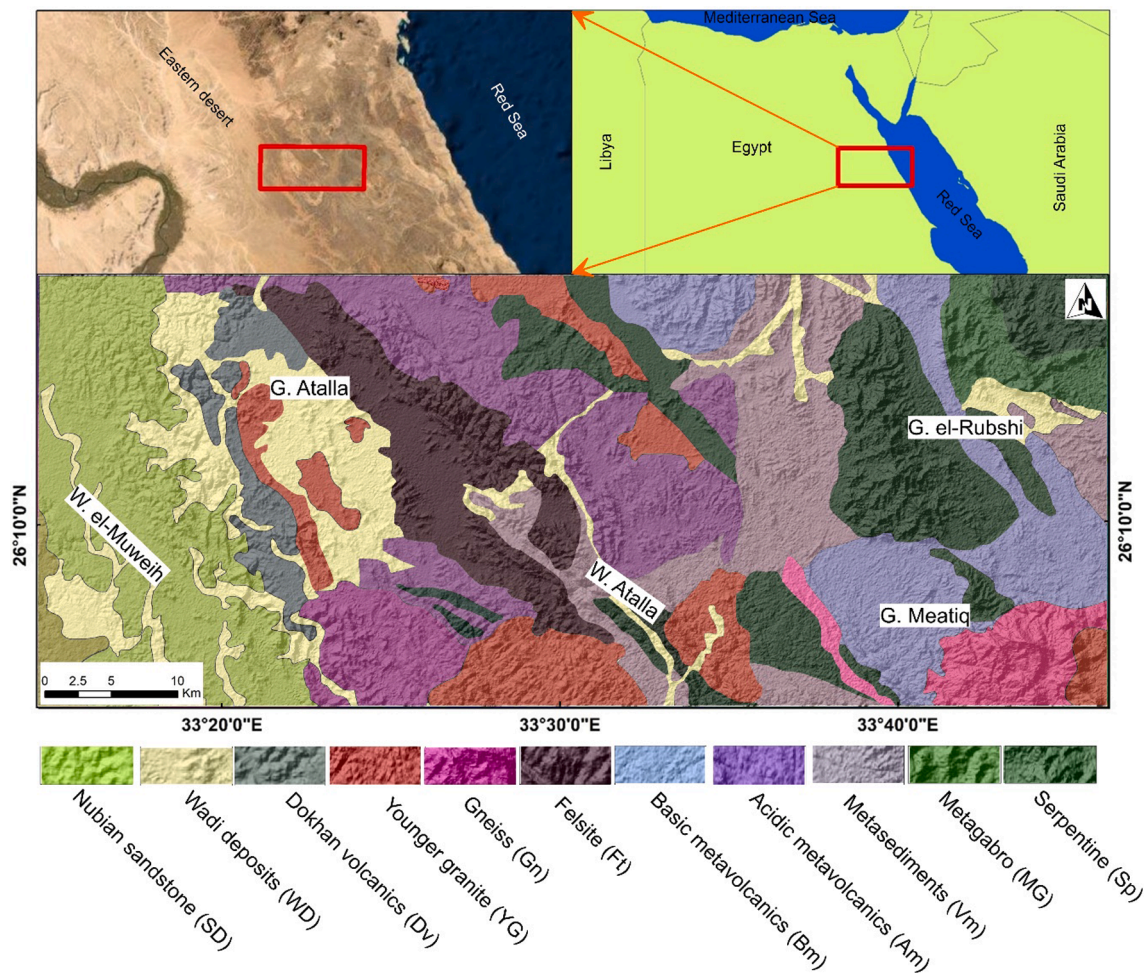


Fig. 1. Location and geological map of the study area (modified after (Conoco, 1987)).

Table 1
Characteristics of Sentinel 2 data.

Band (b)	Central wavelength (μm)	Pixel size (m)
1	0.443	60
2	0.490	10
3	0.560	10
4	0.665	10
5	0.704	20
6	0.740	20
7	0.782	20
8	0.842	10
8a	0.865	20
9	0.945	60
10	1.375	60
11	1.610	20
12	2.190	20

and principal component analysis (Harris, 1989; Harris et al., 1987), and (iii) integration with complementary datasets such as aeromagnetic, satellite imagery, and aerial photographs (Harris et al., 2014; Schetselaar et al., 2000). Although gamma-ray spectrometric datasets are rarely interpreted by themselves, they are effective in mapping surface geology, when they are compared to other airborne geophysical methods (Darnley, 1989). Furthermore, when integrated with other datasets,

considerable results could be attained. Sentinel 2 data as one of the latest optical satellites offering high spectral resolution by 13 channels distributed in Visible Near Infrared (VNIR), and Short-Wave Infrared (SWIR) spectral regions are conjoined and exploited for better mineralogical and lithological recognition (Shebl et al., 2021). The latter is attributed to chemical electronic transitions and vibrations recorded as a radiance then transformed into spectral reflectance uniquely specified to these chemical elements or compounds (Hunt and Ashley, 1979).

Utilizing these data, several MLAs could assign their pixels or sub-pixels to classes in a process known as classification. The latter could be supervised by the analyst who specifies training and testing samples for the classifier in a process called supervised classification. Contrarily, the unsupervised classification process categorizes pixels depending mainly on cluster analysis without analyst interference. Also, classifiers could be widely divided into parametric (preferred with normally distributed data) or non-parametric types. Maximum likelihood classifier is a typical parametric classifier (Grebby et al., 2011) whereas neural networks and support vector machine (SVM) methods are efficient non-parametric classifiers widely applied for lithological mapping using remote sensing data (Liesenberg and Gloaguen, 2012; Pal and Mather, 2005). Support vector machine, as a well-known efficient classifier in performing lithological mapping (has higher ability in separating complicated boundaries) utilizing geological and geophysical datasets (Bachri et al., 2019; Mou et al., 2015; Shebl and Csámer, 2021), was

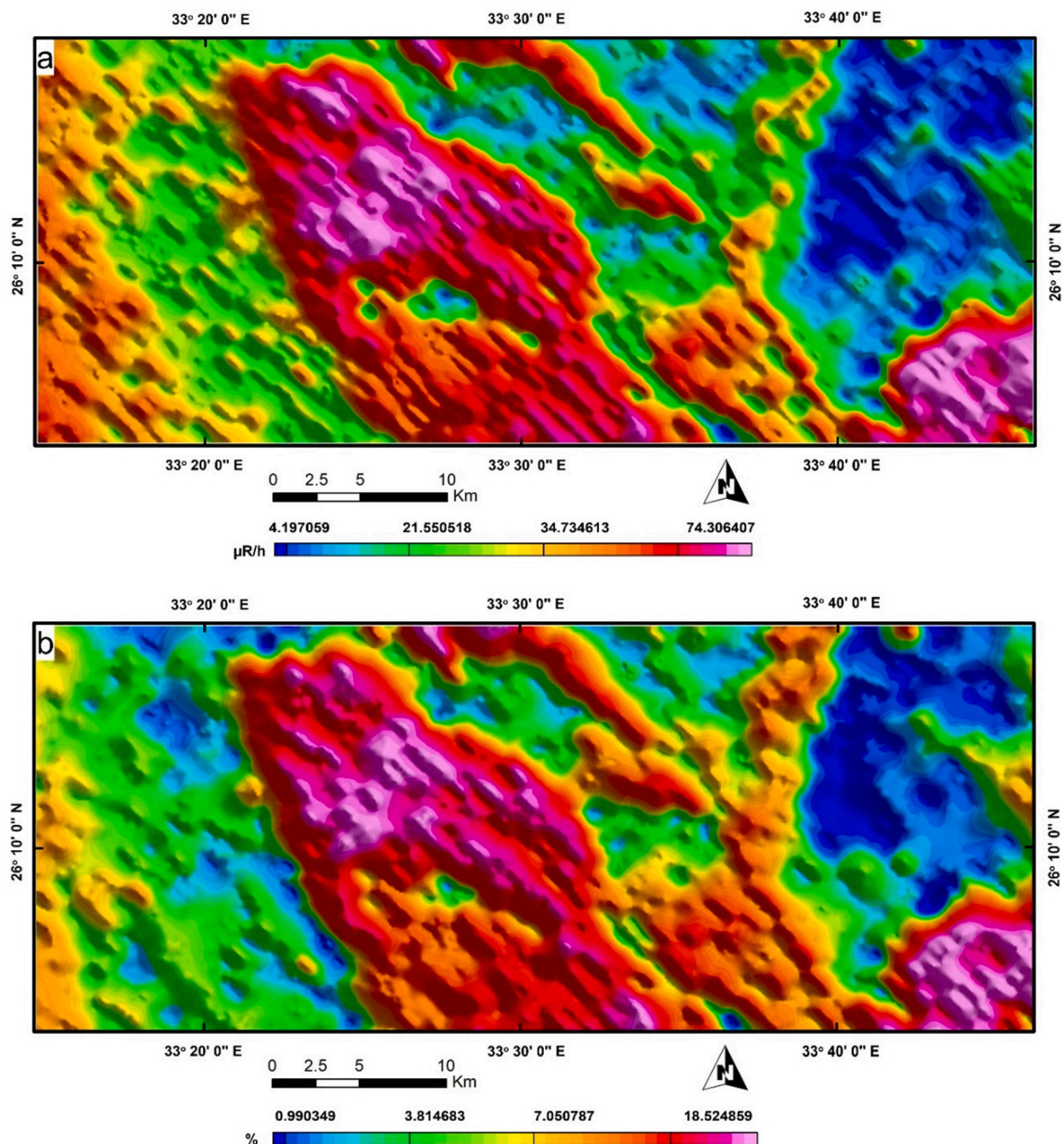


Fig. 2. Visual representation of (a) the total count, and (b) potassium gamma-ray spectrometric maps.

implemented in the current study.

The current study aims to use the power of SVM to classify the fused remote sensing (Sentinel 2 data) and radiometric measurements to enhance lithological differentiation over a case study from Arabian Nubian Shield (well-known for the abundance of various mineral deposits). Thus, orogenic belts could be efficiently mapped, or at least the old geological maps could be renovated by adopting this approach.

2. Study area and geological setting

The investigated area (Atalla area and its environs) is located between latitudes 26° 05'' to 26° 17'' N and longitudes 33° 15'' to 33° 46'' E

as shown in Fig. 1. This area is covered by a part of Meatiq Group gneiss (one of the oldest Precambrian rock units in the Nubian shield), ophiolitic constituents (serpentinite, metagabbro, and amphibolite within volcanoclastic metasediments, as a matrix of the ophiolitic mélange), metavolcanics (acidic and basic), syn- to late-tectonic granitoids (Atalla granite), granodiorites, felsites, post tectonic granitoids, Nubian sandstone and wadi deposits (ABD EL MONSEF, 2020; Conoco, 1987; Hamimi et al., 2020; Shebl et al., 2021). Thus, 13 lithological targets representing igneous, metamorphic, and sedimentary rocks are included. They exemplify most of the major lithological units within the Nubian shield. The study area includes several alteration zones around the Atalla gold mine at the central part of the investigated area. The

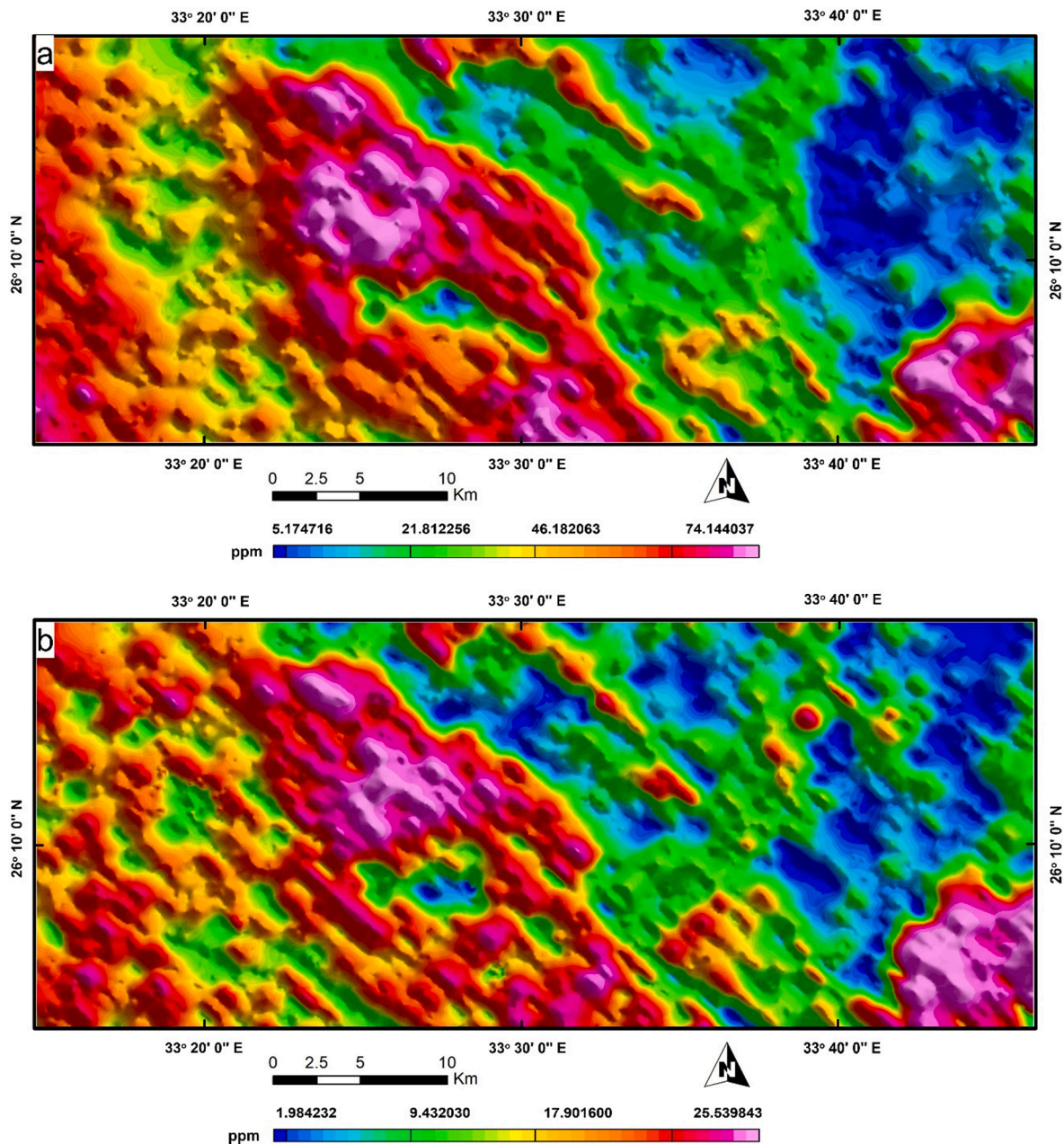


Fig. 3. Visual representation of (a) thorium, and (b) uranium gamma-ray spectrometric maps.

recently published geological maps (ABD EL MONSEF, 2020; Hamimi et al., 2020; Shebl et al., 2021) were built upon field check and petrographic studies, and cover significant parts of the study area, however, CONOCO (Continental Oil Company) geological map (Conoco, 1987) covers the whole area as shown in Fig. 1.

The four geological maps besides remote sensing data were integrated to build a reference geological map suitable to specify training and testing samples (depending mainly on radiometric and spectral signatures). The availability of these maps, reasonable representation of various rock units, and finally, the area accessibility for executing a reconnaissance field check strongly recommended this case study for applying the current method.

3. Materials and methods

3.1. Sentinel 2 (S2)

Sentinel-2 satellites (S2A and S2B) are provided with Multi-Spectral Instrument (MSI) to image the earth at the visible and broad NIR, red edge, narrow NIR, and SWIR regions from an altitude of approximately 786 km as shown in Table 1. MSI implies a push-broom imaging concept and has a swath width of 290 km (20.6° field of view). For the current study, S2A_MSIL1C_20201231T082341_N0209_R121_T36RWP_20201231-T100300.SAFE Top-Of-Atmosphere (TOA) scene was accessed via

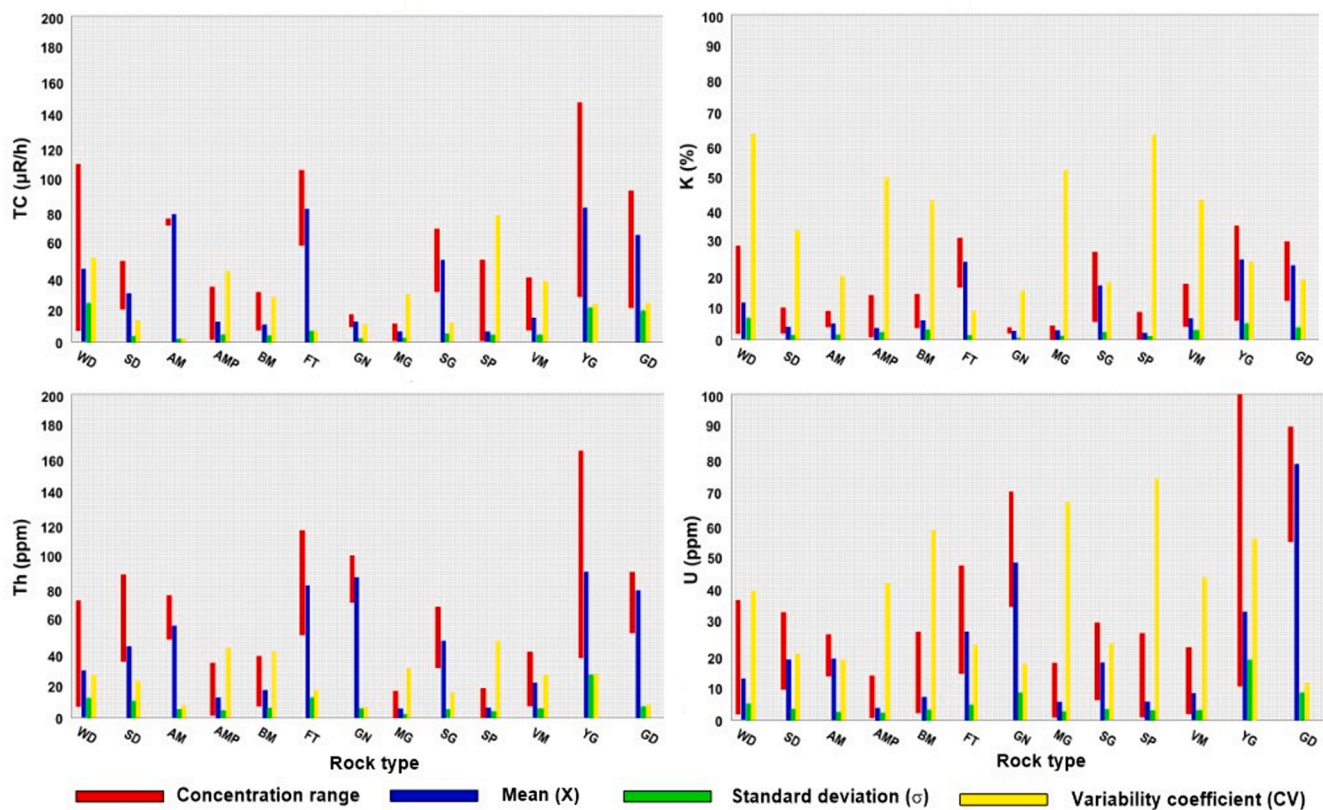


Fig. 4. Radioelements statistical analysis of the investigated rock units.

European Space Agency (ESA) Open Access Hub then atmospherically corrected using Sen2Cor (S2C) Level-2A offline processor to get S2A_MSI-L2A_20201231T082341_N9999_R121_T36RWP_20210314T120209. SAFE Bottom-Of-Atmosphere (BOA) reflectance scene, which is handled using Sentinel Application Platform (SNAP). Then VNIR and SWIR bands are resampled to a pixel size of 10 m. All 60 m pixel-sized S2A_MSI_L2A bands were excluded from this analysis due to their smaller spatial resolution.

3.2. Gamma-ray and fused data

The gamma-ray spectrometric data were obtained through (Aero-Service, 1984). The data are available in the form of contour maps. For further processing, the maps have been digitized and then the resultant data points (29069, 9334, 7382, and 10,216 for the total count, potassium, thorium, and uranium respectively) are visually displayed in Figs. 2, and 3. Also, the raw data have been statistically analyzed to determine minima, maxima, means (X), standard deviations (σ), and coefficient of variability (CV%) of every variable as shown in Fig. 4. It is worth mentioning that if the CV% is less than 100 %, the variables tend to exhibit normal distribution based on the relation of $CV = (\sigma/X) \times 100$.

Then, the projections of these radiometric points are transformed to WGS_1984 UTM_Zone_36N. Thereafter, a spatial interpolation using inverse distance weighted method (IDW) was carried out to generate four rasters of K, Th, U, and their sum (Figs. 5 and 6) having the same spatial size (10 m) and dimensions in pixels (5338×2156) of Sentinel 2. IDW method assigns the predicted cell value using a weighted

combination of surrounding sample points, and as the name suggests the predicted weight is a function of the inverse distance (Eckstein, 1989). IDW method was adopted in this study due to its wide usage, and its convenience with our data nature (densely distributed points). To form extended pixel vectors by aggregating spectral (Sentinel 2) and radiometric (gamma-ray), the data is fused using the stacked-vector method (Richards and Jia, 1999; Shebl and Csámer, 2021) to build 5 main fused classifier inputs (S2, S2 + K, S2 + Th, S2 + U, S2 + TC) having the same dimensions and pixel sizes.

4. Reference geological map and sampling

An efficient reference geological map is crucial for influential sampling which has a pivotal role in enhancing classification accuracy (Kumar et al., 2020). To be away from scale inconveniences accompanied with traditional geological maps, we built a high-resolution geological map (10 m) through enhancing the spectral separability between the lithological targets by integrating four geological maps (as they covered different parts of the study area in varied mapping scales, that could help in better interpretation compared to single lithologic map), with False Color composites (FCCs), Principal Component Analysis (PCA), and Independent Component Analysis (ICA). Each technique accentuated certain lithological and structural features as shown in Figs. 7 and 8, however, it was found that FCC 12/6/2 RGB gives reasonable demarcation for all the required targets (that could be better resolved with reference to PCA and ICA results), hence it is implemented as the input for feature selection and sampling.

FCC 12/6/2 RGB (10 m reference map) displayed a reasonable

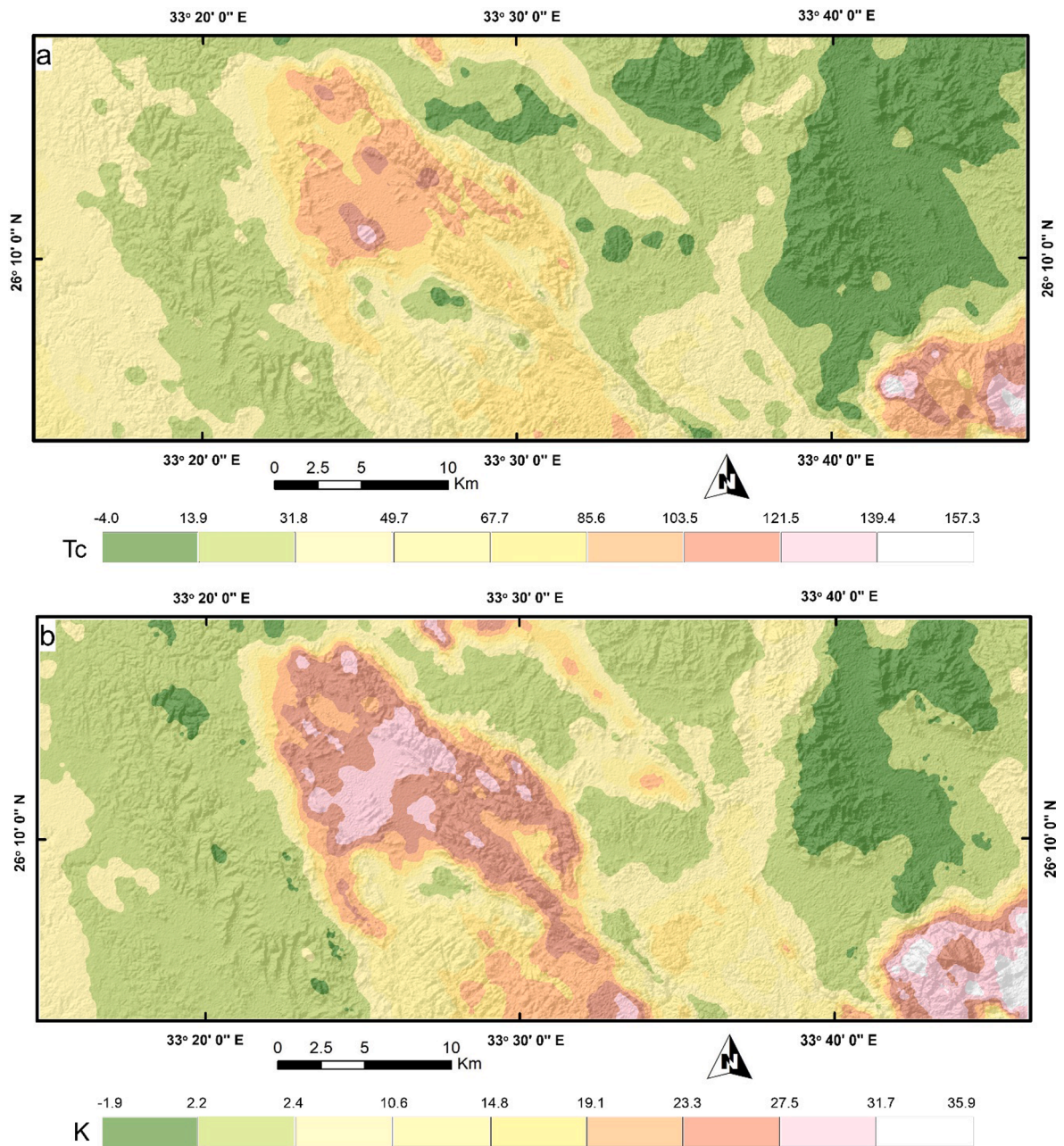


Fig. 5. IDW results showing (a) total count, and (b) potassium gamma-ray spectrometric maps.

separation for each lithological target, thus a random sampling method was espoused to select training samples from 13 classes. The lithological classes are Gneiss (Gn), Serpentinities (Sp), Metagabbros (MG), Amphibolites (Amp), Volcaniclastic metasediments (VM), Acidic metavolcanics (AM), Basic metavolcanics (BM), Syn to late tectonic granites (SG), Granodiorite (Gd), Felsites (Ft), Younger Granites (YG), Nubian Sandstone (SD), and Wadi Deposits (WD). Researchers always recommend 70%-80% to 30%-20% percentage of training to testing samples respectively (Kumar et al., 2020). Thus, area-wise training samples with percentages ranging between 70 and 80% were extracted depending on the spectral characteristics of each rock unit (Table 2). Training samples are utilized to build spectral reflectance curves for the utilized samples and the total reflectance values as shown in Fig. 9, which reveals how

close are the rock units' reflectances. Training areas separability was assessed via Jeffries-Matusita (JM) distance which represents dissimilarity between 2 different classes on a scale ranging from 0 to 2 (Kumar et al., 2020) as displayed in Table 3. Testing data has never been seen by the classifier and carefully selected to evaluate the classifier performance.

4.1. Support vector Machine

A widely used classifier that performs optimal separation by delineating the closest training samples in the spectral space (Pal and Mather, 2005; Scholkopf and Smola, 2018). SVM is originally a binary classifier that tries to separate two classes via building a decision boundary

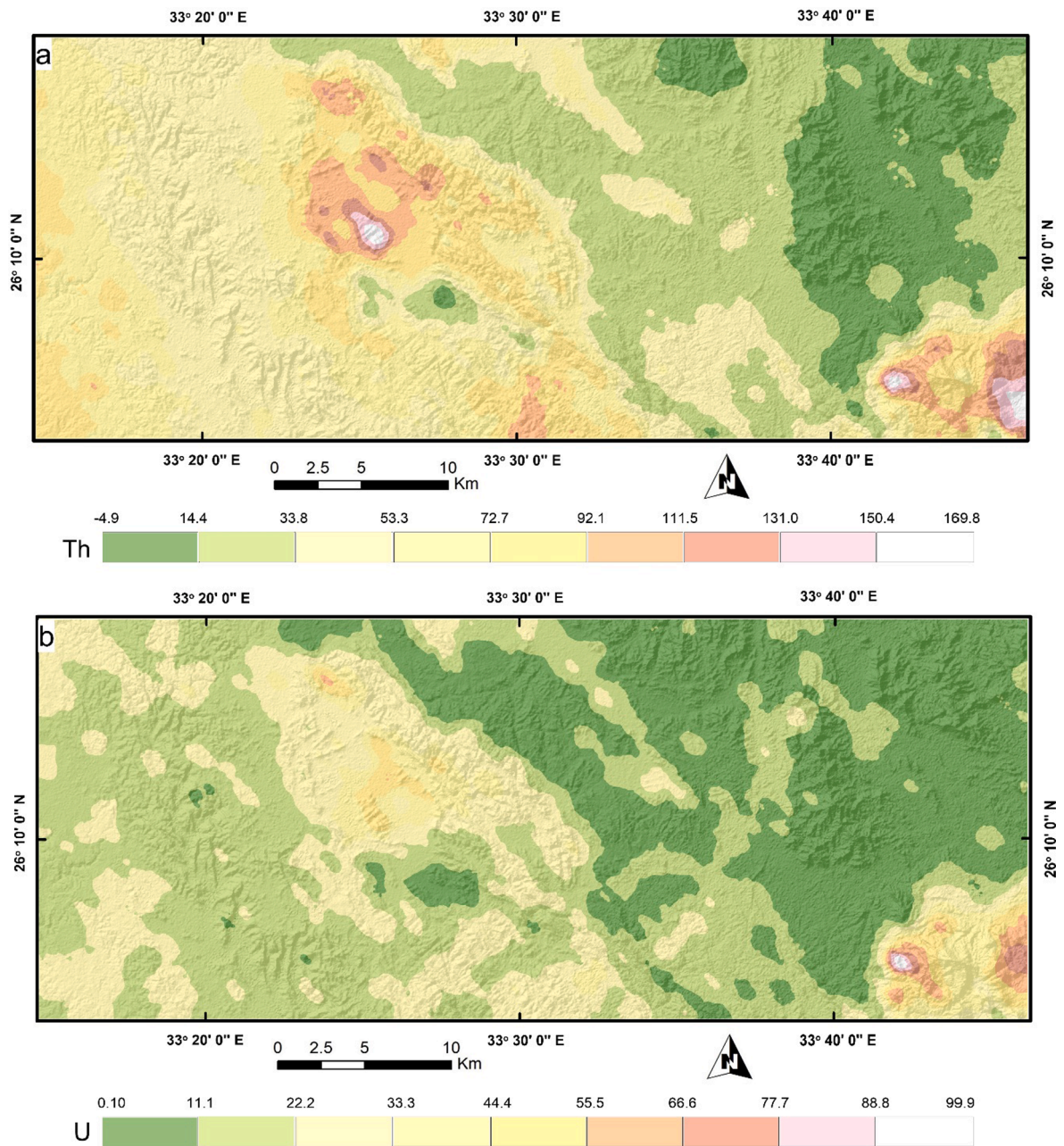


Fig. 6. IDW results showing (a) thorium, and (b) uranium gamma-ray spectrometric maps.

between them, then several classes could be distinguished via a hyper-plane by repeating the process between classes (Maxwell et al., 2018). In this study, Radial Basis Function (RBF) was preferred as a kernel function as suggested by (Bachri et al., 2019). Gamma in kernel function was selected to be the inverted number of the participated bands, a penalty parameter for misclassifications was set to 100 to ensure reliable results. After this optimization process, the five inputs (S2, S2 + K, S2 + Th, S2 + U, S2 + TC) were classified using the same training samples via SVM classifier.

5. Results and discussion

Five classification processes were executed via SVM using the same training areas to ensure wise comparison. Accuracies were assessed via confusion matrix, User Accuracy (UA), Producer accuracy (PA), Overall Accuracy (OA), and Kappa Coefficient (K). Generally, blending of S2 optical data with any element of the implemented spectrometric data enhances the allocation OAs, whereas S2 only delivers the least OA of 77.93% compared to 82.07%, 83.04%, 84.76%, 85.70% using S2 + U,

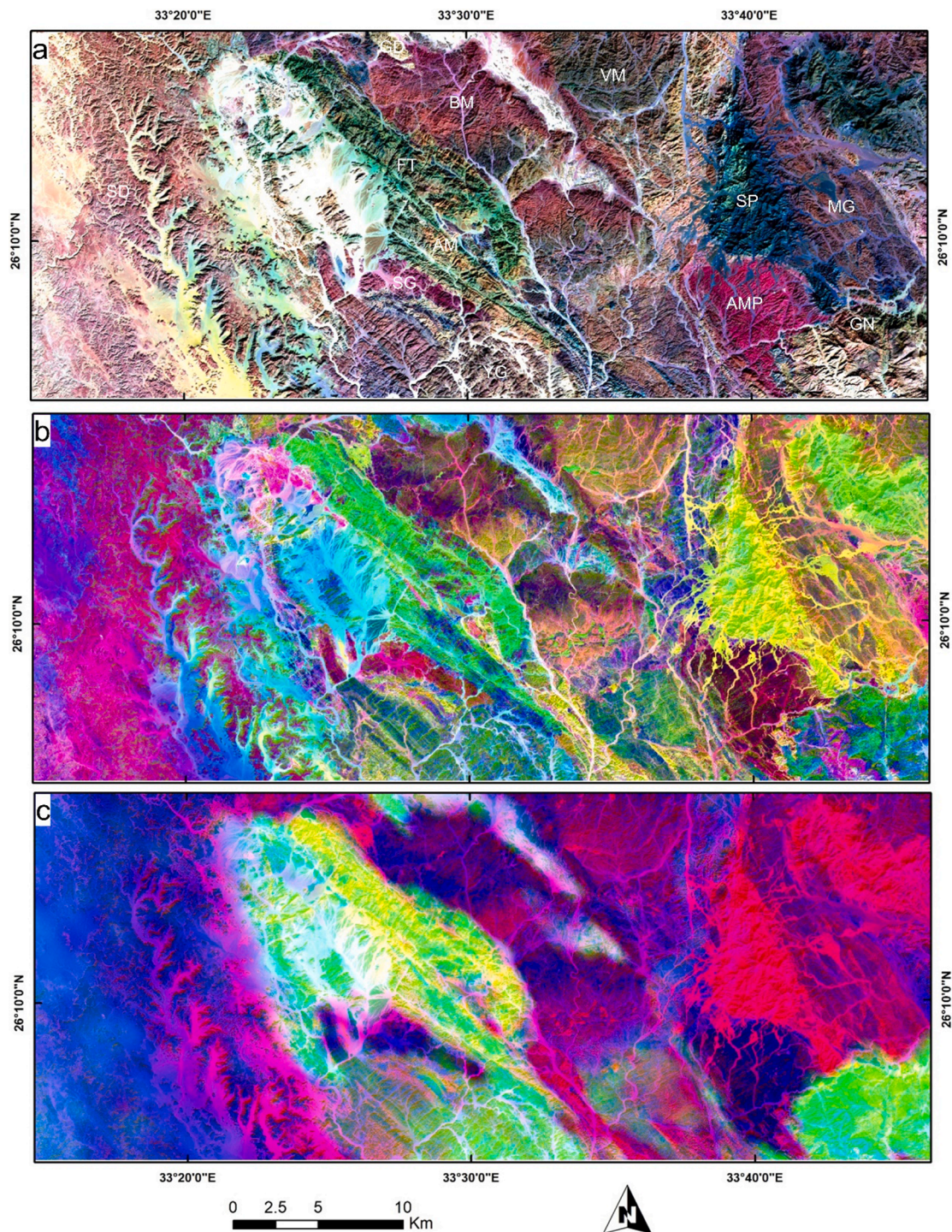


Fig. 7. Lithological discrimination via (a) FCC 12/6/2 (b) S2 data PC5, PC2, and PC1 (c) fused S2 and total count PC3, PC1, and PC2, in RGB respectively.

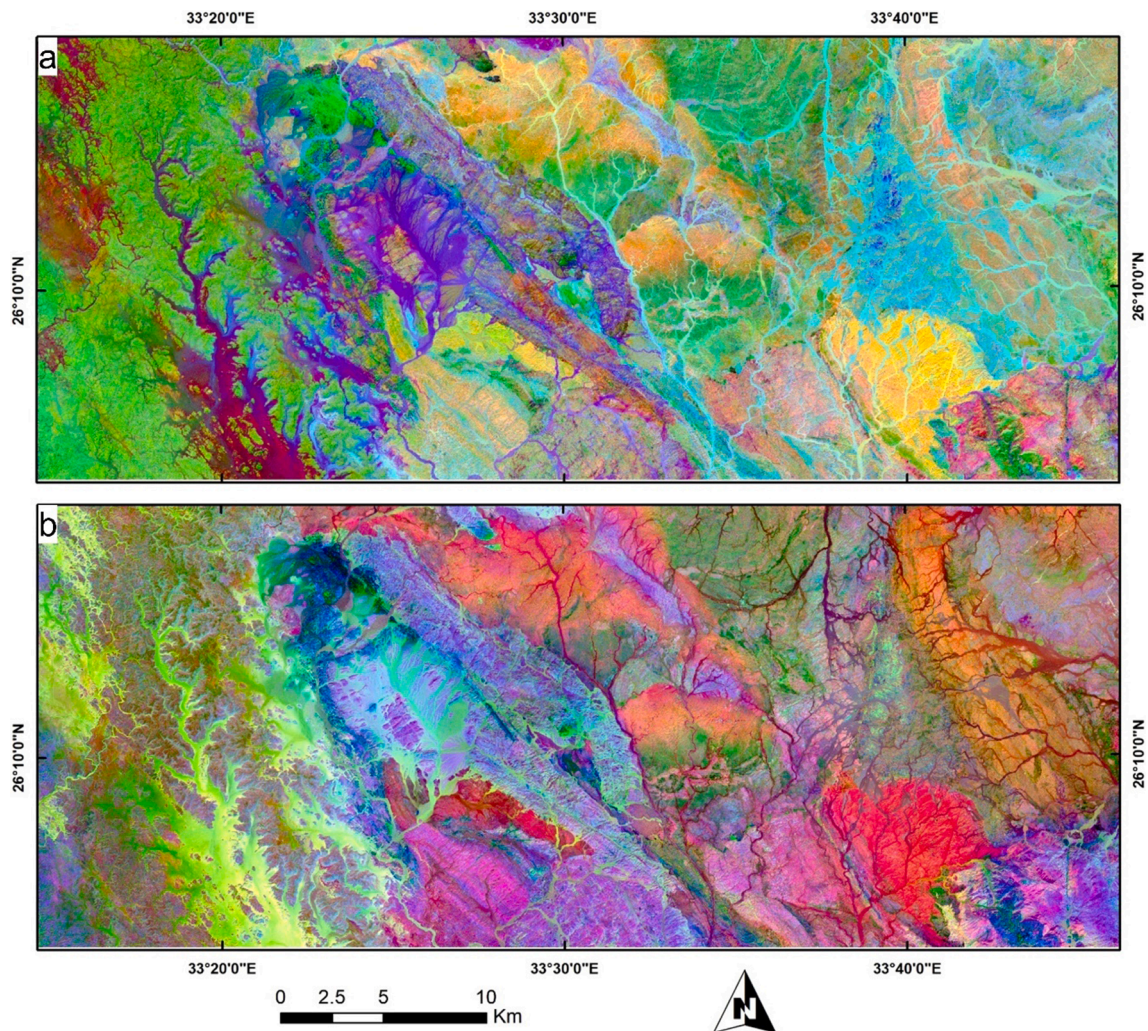


Fig. 8. Lithological discrimination via, (a) S2 data IC2, IC4, and IC1, (b) fused S2 and potassium data IC1, IC2, and IC3, in RGB respectively.

Table 2
Characteristics of training and testing data.

Classes	Training data		Testing data		total
	pixels	%	pixels	%	
Sp	3098	77.6	891	22.4	3989
Amp	2029	78.4	557	21.6	2586
VM	3204	76.2	996	23.8	4200
SG	2016	79	531	21	2547
Ft	3037	78.9	808	21.1	3845
YG	3095	78.4	852	21.6	3947
Gn	1580	75	525	25	2105
MG	2204	77.2	648	22.8	2852
WD	3386	71.3	1357	28.7	4743
SD	2654	79	703	21	3357
GD	955	80	233	20	1188
AM	1993	74.1	693	25.9	2686
BM	3396	79	902	21	4298

S2 + K, S2 + Th, and S2 + Tc, respectively. Variations in OAs are attributed to different concentrations of K, Th, and U in the lithological classes, as displayed in Fig. 4. The final thematic layers are shown in Figs. 10 and 11 to closely compare the separability of the lithological targets depending on the classifier feed. For instance, at the western side

of the study area, where the basement rocks are unconformably overlain by sedimentary rocks and through the drainage area (wide wadis), we can see how accurate the S2 + total count map is, in diagnosing not only the wadi deposits but also could specifying other lithological compositions within it as represented by serpentinite apparition. The presence of serpentinite along these wadis may be attributed to the accumulation of weathered serpentinite fragments in low topographic regions. In the eastern part of the maps, a massive serpentinite block (Gabal El-Rubbshi) is clearly isolated in all the maps however, if we intensively focused on the rock unit, we could notice the random distribution of error pixels (light cyan) within it in case of using S2 only as a classifier input. These misclassified pixels are almost well-classified by adding radiometric data. This in turn confirms the powerful function of the integrated datasets in lithological allocation and gives a possible way to weed out the salt and pepper effect that is predominantly associated with lithological classifications. This is meticulously confirmed over the resultant maps, for example, investigating the inconspicuous picking out of amphibolite block (blue color) using S2 compared to pure identification when Th, K, U, or Tc data are added to S2. Metagabbroic rocks at the eastern part of the map confirmed that the area of the mapped rock unit is affected by the utilized data, where significant parts are misclassified as volcanoclastic metasediments using only spectral S2 or S2 +

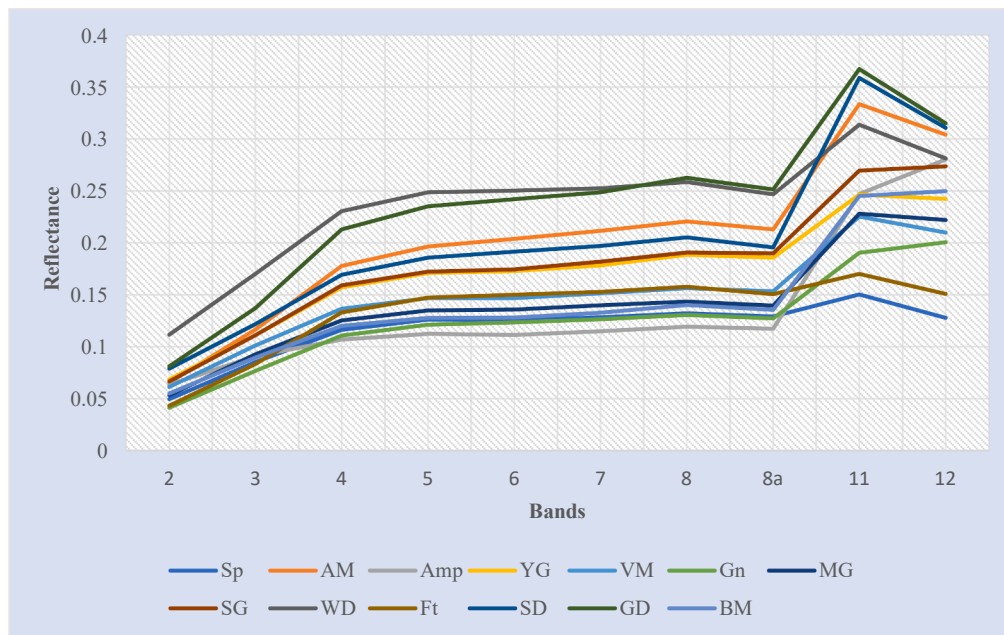


Fig. 9. Average spectral reflectances of the classified lithological classes.

Table 3

Average Jeffries–Matusita distances respective to Sentinel 2 and Sentinel 2 + total count images.

S2	Amp	VM	BM	SG	Ft	AM	YG	Gn	MG	WD	SD	GD
Sp	1.998616	1.776571	1.992235	1.989125	1.573896	1.966661	1.857694	1.823067	1.975211	1.895772	1.999251	1.957564
Amp	0	1.987557	1.802624	1.996974	2	2	1.996036	1.9516	1.986816	1.999694	1.998009	2
VM	0	0	1.886733	1.624341	1.982777	1.921096	1.591896	1.44851	1.449328	1.918179	1.962584	1.892818
BM	0	0	0	1.927016	1.998272	1.977543	1.97202	1.778091	1.567768	1.999146	1.971037	1.985086
SG	0	0	0	0	1.998915	1.759178	1.315719	1.515118	1.742344	1.996513	1.997419	1.917847
Ft	0	0	0	0	0	1.981898	1.928999	1.874673	1.998074	1.985269	1.999999	1.967379
AM	0	0	0	0	0	0	1.80947	1.739928	1.967141	1.994842	1.97228	1.481924
YG	0	0	0	0	0	0	0	1.379709	1.914998	1.985249	1.999305	1.788691
Gn	0	0	0	0	0	0	0	0	1.720845	1.98431	1.990948	1.749116
MG	0	0	0	0	0	0	0	0	0	1.986008	1.98783	1.974849
WD	0	0	0	0	0	0	0	0	0	0	1.989247	1.986303
SD	0	0	0	0	0	0	0	0	0	0	0	1.958738
S2 + Tc	Amp	VM	BM	SG	Ft	AM	YG	Gn	MG	WD	SD	GD
Sp	2	1.839695	1.994198	1.994586	1.999498	1.999484	1.979653	1.999986	1.985973	1.92911	1.99948	1.996994
Amp	0	1.997435	1.995261	2	2	2	2	2	1.993575	2	1.999967	2
VM	0	0	1.924475	1.840753	1.999884	1.999939	1.955283	1.999996	1.855786	1.951438	1.984327	1.996679
BM	0	0	0	1.995896	2	2	2	2	1.811586	2	1.989866	2
SG	0	0	0	0	1.999716	1.976653	1.896835	1.999477	1.997263	1.998281	1.99955	1.982959
Ft	0	0	0	0	0	1.981588	1.940945	1.910365	2	1.99677	2	1.966362
AM	0	0	0	0	0	0	1.906471	1.912597	2	1.998681	1.999995	1.600687
YG	0	0	0	0	0	0	0	1.686131	2	1.995945	2	1.853869
Gn	0	0	0	0	0	0	0	0	1.720845	1.98431	1.990948	1.749116
MG	0	0	0	0	0	0	0	0	0	1.986008	1.98783	1.974849
WD	0	0	0	0	0	0	0	0	0	0	1.989247	1.986303
SD	0	0	0	0	0	0	0	0	0	0	0	1.958738

U content, however, an outright identification is noticed for S2 + K, S2 + Th, and S2 + Tc with slight variances between them corresponding to their content as shown in Fig. 4. All these differences are attributed to the varied radiometric contents of each rock unit as shown in Fig. 4 and the consequent algorithm response to these inputs. Generally, increasing the operative amount of data (included in the classification process) enhances the allocation process (Shebl and Csámer, 2021) as observed for the S2 + total count map.

Statistically, the sublimity of the additional spectrometric data besides S2 multispectral bands is explicitly distinct when comparing the JM distance matrix of the former to the latter, as displayed in Table 3. S2 + TC JM matrix values are optimum (2) or very close to 2 for the most lithological classes. That differs from the S2 JM dissimilarity matrix, where several values are far from the optimum separation. Pair separation comparison is carried out to closely monitor the least separation values or classes displaying higher similarities. For S2 data solely,

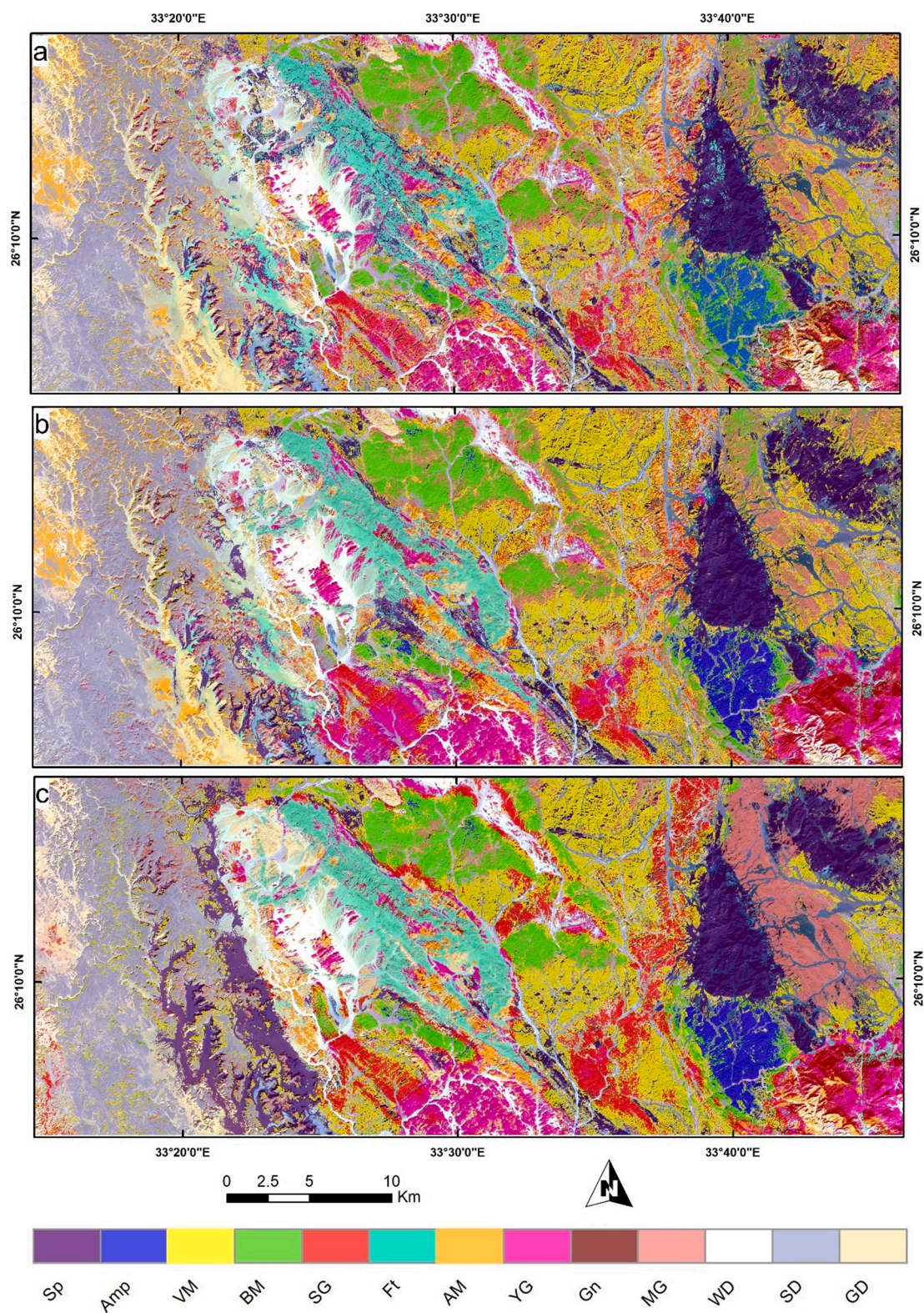


Fig. 10. Lithological thematic maps classified using SVM based on (a) Sentinel 2, (b) Sentinel 2 and Uranium, (c) Sentinel 2 and Potassium.

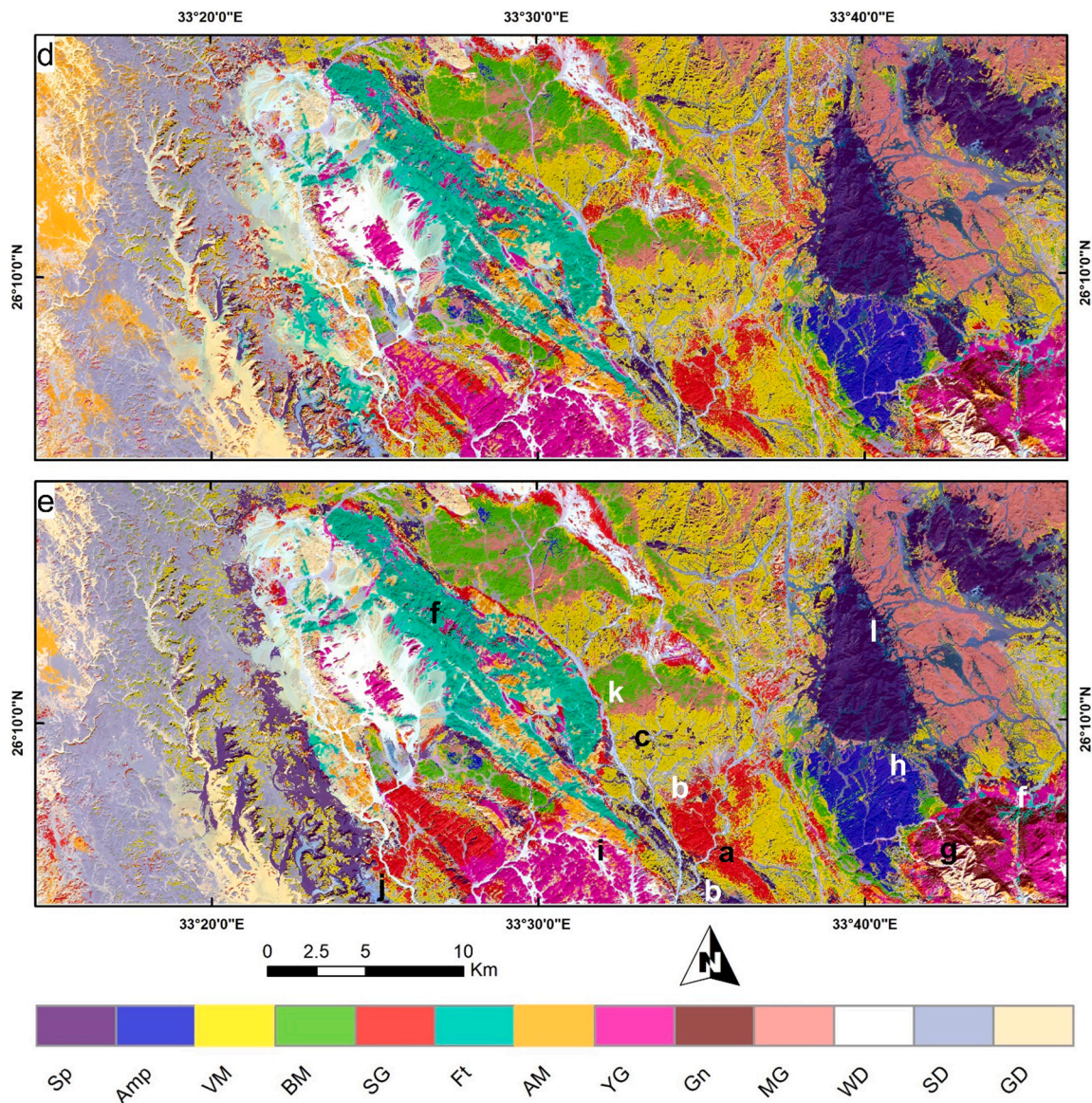


Fig. 11. Lithological thematic maps classified using SVM based on (a) Sentinel 2 and thorium, (b) Sentinel 2 and total count (letters are related to the approximate locations of field photographs (Fig. 13) and names of rock units).

classes that offer higher similarities are SG and YG, YG and Gn, VM and Gn, VM and MG, AM and GD, SG and Gn, BM and MG, Sp and Ft, VM and YG, VM and SG, by JM distances of 1.315, 1.379, 1.44, 1.44, 1.48, 1.515, 1.56, 1.57, 1.59, and 1.62, respectively. These pairs are more distinct and separable using S2 + Tc data by JM values of 1.89, 1.68, 1.85, 1.99, 1.60, 1.99, 1.85, 1.95, and 1.84. This in turn shows how well the added spectrometric total count layer helps in the discrimination spectrally similar classes and thus delivering higher OAs.

A closer investigation of each rock unit through confusion matrix, UA, PA reveals that spectral confusion is always evident with volcanoclastic metasediments (VM), granitic rocks (SG, YG, GD), and gneisses, as shown in Fig. 12. For VM and as a matrix for ophiolitic mélangé, it is expected to find several inclusions of ophiolitic blocks which, in turn, increases intra-class variability and leading to faint separation, as shown

in Table 4. Similarly, SG, YG, GD, and Gn have very similar mineralogical and chemical compositions, thus the classifier sometimes is confused and may classify SG as YG or Gn, and vice versa (Commission and Omission errors) dropping down UA and PA for these classes as confirmed via the JM matrix, where the least separation is almost accompanied with one of these classes. On the other side, Sp, Ft, SD, WD, Amp, and MG are well separated and resulted in higher PA, UA, and JM distances due to reasonable discrimination from the rest of the classes almost with the five inputs introduced to SVM.

Slight inequalities in UA, PA, and JM values for the same class are attributed to varied powers of the input data in distinguishing such classes. For instance, the spectral remarkability of Sp with S2 data, depletion in U, K, and Th (Fig. 4) results in a low total count makes Sp easily separated from other rock units (PA is over 90% for all the inputs).

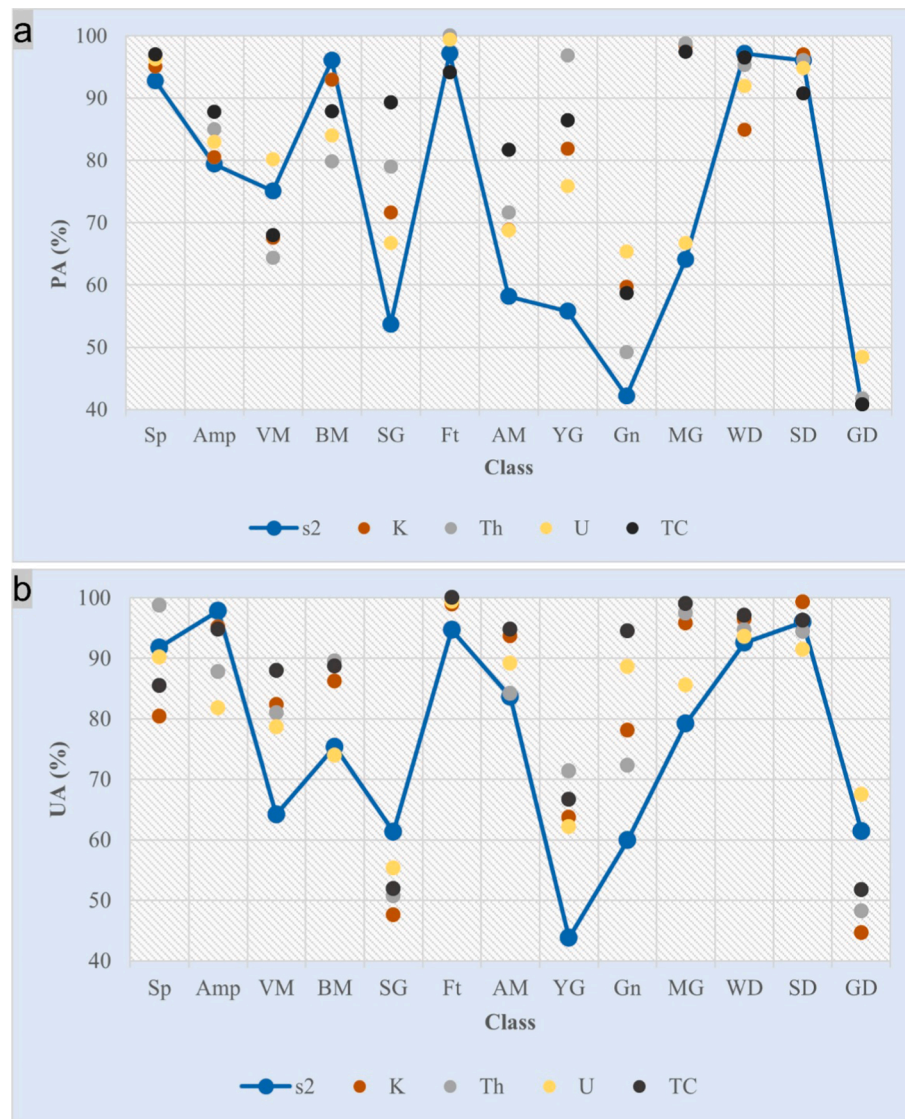


Fig. 12. (a) Producer accuracy (PA), and (b) User accuracy (UA) for each class showing how well the classification is enhanced by adding any of K, U, Th, or Tc to Sentinel 2 data (solid line). This is obviously manifested just by comparing the number of points above and below the solid line.

Similarly, felsites with higher content of spectrometric data and distinguished spectral signature are rarely misclassified (PA is over 90% for all the inputs).

Contrarily, the K content of VM is very close to that of AM and BM, VM has low U content closer to SP and AM content, as well as the mixed spectral signatures included from predominant ophiolitic blocks and dykes resulting in PA less than 80 % for all the data inputs. In this way, a rock unit is distinguished from the other by its K, Th, U, or total content added to its spectral reflectance. It is worth mentioning that K, Th, U, and Tc spectrometric maps poorly differentiate all the rock units, as shown in spectrometric data maps, however, fusing any of these maps with S2 multispectral data boost the classification process and separation accuracy as shown in Fig. 11b (great representation compared to reference geological map) and proved in Table 3 and Table 4. It was confirmed by Fig. 12, which shows the effect of the added radiometric data compared to Sentinel 2, depending on the lithological target composition.

Besides visual comparison with reference and previous geological maps, the statistical accuracy assessment of the thematic results, a field check was performed to validate our results. Fieldwork was carried out by investigating the mapped rock units (Fig. 13) via 30 randomly distributed stations across the area. A great coincidence was evident among most of the investigated rocks, confirming the validity of the utilized approach in an efficient lithologic mapping.

Through the results of the current study, we strongly recommend incorporating gamma-ray data in predictive lithological mapping utilizing MLAs. The study highlighted the effectiveness of the approach in decreasing fuzzy representation of the information classes and increasing the cohesion and accuracy of the categorization process. Most prior studies implemented multispectral data in geological mapping, however, augmentation with gamma-ray data generates improved thematic maps. The best results are obtained when all the chemical concentrations (Total count) are utilized. SVM proved its efficiency in the reasonable prediction of the lithological classes. In our opinion, this

Table 4

Error matrix, Producer Accuracy (PA), User Accuracy (UA), Overall accuracy (OA), and Kappa coefficient (K) of SVM classification results.

S2	Sp	Amp	VM	BM	SG	Ft	AM	YG	Gn	MG	WD	SD	GD	Tot	PA	UA
Sp	826	0	13	0	0	23	1	1	13	0	21	0	2	900	92.7	91.78
Amp	0	442	0	10	0	0	0	0	0	0	0	0	0	452	79.35	97.79
VM	16	0	725	6	63	0	15	250	21	34	0	0	0	1130	75.05	64.16
BM	0	115	0	866	0	0	1	0	7	155	0	0	5	1149	96.01	75.37
SG	0	0	56	0	285	0	60	33	20	3	0	0	8	465	53.67	61.29
Ft	36	0	0	0	0	785	0	0	0	0	8	0	0	829	97.15	94.69
AM	0	0	1	0	0	0	403	27	20	0	0	23	8	482	58.15	83.61
YG	0	0	149	0	143	0	86	475	182	2	0	0	47	1084	55.75	43.82
Gn	1	0	1	2	23	0	18	13	221	26	0	0	64	369	42.1	59.89
MG	0	0	10	14	14	0	17	26	26	415	0	2	0	524	64.04	79.2
WD	12	0	11	3	2	0	35	23	15	3	1318	2	0	1424	97.13	92.56
SD	0	0	0	1	1	0	6	0	0	10	10	675	0	703	96.02	96.02
GD	0	0	0	0	0	0	51	4	0	0	0	1	89	145	39.91	61.38
Tot	891	557	966	902	531	808	693	852	525	648	1357	703	223	9656	OA = 77.93	K = 0.757
S2 + Tc	Sp	Amp	VM	BM	SG	Ft	AM	YG	Gn	MG	WD	SD	GD	Tot	PA	UA
Sp	864	0	5	0	0	40	0	0	0	0	0	102	0	1011	96.97	85.46
Amp	0	489	0	26	0	0	0	0	0	1	0	0	0	516	87.79	94.77
VM	15	0	656	17	54	0	0	0	0	0	2	2	0	746	67.91	87.94
BM	0	68	0	792	0	0	0	0	0	14	19	0	0	893	87.8	88.69
SG	0	0	300	56	474	6	0	73	0	0	4	0	0	913	89.27	51.92
Ft	0	0	0	0	0	760	0	0	0	0	0	0	0	760	94.06	100
AM	0	0	0	0	0	0	566	13	0	0	0	0	18	597	81.67	94.81
YG	0	0	0	0	0	0	39	736	217	0	0	0	112	1104	86.38	66.67
Gn	0	0	0	0	0	0	0	16	308	0	0	0	2	326	58.67	94.48
MG	1	0	0	5	0	0	0	0	0	631	0	0	0	637	97.38	99.06
WD	0	0	0	4	0	0	0	0	0	0	678	16	0	698	96.44	97.13
SD	11	0	5	2	3	0	14	11	0	2	0	1231	0	1279	90.71	96.25
GD	0	0	0	0	0	2	74	3	0	0	0	6	91	176	40.81	51.7
Tot	891	557	966	902	531	808	693	852	525	648	703	1357	223	9656	OA = 85.70	K = 0.843
S2+U	PA		UA		S2+K		PA		UA		S2+Th		PA		UA	
Sp	96.18		90.21		Sp	95.06		80.36		Sp	96.97		98.74			
Amp	82.94		81.77		Amp	80.43		95.12		Amp	84.92		87.76			
VM	80.12		78.66		VM	67.49		82.32		VM	64.29		80.96			
BM	83.92		73.93		BM	92.9		86.21		BM	79.82		89.55			
SG	66.67		55.31		SG	71.56		47.56		SG	78.91		50.67			
Ft	99.38		99.38		Ft	99.63		98.89		Ft	100		99.63			
AM	68.69		89.14		AM	68.83		93.71		AM	71.57		84.21			
YG	75.82		62.18		YG	81.81		63.71		YG	96.83		71.37			
Gn	65.33		88.63		Gn	59.62		78.05		Gn	49.14		72.27			
MG	66.67		85.54		MG	98.3		95.79		MG	98.77		97.41			
WD	91.89		93.62		WD	84.89		96.48		WD	95.28		94.66			
SD	94.74		91.48		SD	97.01		99.27		SD	96.02		94.41			
GD	48.43		67.5		GD	40.81		44.61		GD	41.7		48.19			
Tot	OA=82.07		K=0.803		Tot	OA=83.04		K=0.814		Tot	OA=84.76		K=0.833			

method is amenable to a wide spectrum of lithological conditions as it delivers acceptable results in complicated basement rocks of highly similar spectral signatures. Our future research will be executed over vegetated geological terranes aiming to enhance the allocation process of rocks suppressed by vegetation cover using radiometric data and the power of deep learning algorithms.

6. Conclusion

The study proposed an approach for efficient lithological mapping through applying Support Vector Machine (SVM) to integrated gamma-ray spectrometric data with Sentinel 2 data (high spatial and spectral resolutions) leading to monitoring minute variations among 13 classes and discrimination of closely related spectral signatures. The study concludes the following points:

- 1- FCC, PCA, and ICA conjugated with geological maps could generate a reference lithological map (with the same scale and pixel size of classification inputs) that is more compatible for training and testing samples extraction compared to conventional maps.
- 2- Gamma-ray spectrometric data delivers a poor lithological classification and cannot separate rock units compared to Sentinel 2 data. Fusing S2 data with K, Th, U, or their sum results in better lithological identification. The overall classification accuracies for Sentinel 2, S2 + U, S2 + K, S2 + Th, and S2 + TC are 76.5%, 77.93%, 82.07%, 83.04%, 84.76%, and 85.70%, respectively. The optimum result is achieved by combining S2 bands with total count gamma-ray data.
- 3- The study strongly recommends blending the power of MLAs with the combined S2 and gamma-ray data in further geological mapping or in renovating older geological maps, especially at terranes that are



Fig. 13. Field photographs showing the mapped rock units of (a) Atalla granite, (b) Sharp contact between Atalla granite and metavolcanics, (c) Volcaniclastic metasediments, (d) Massive serpentinite blocks transformed into talc carbonates along zones of thrusting, (e) Meatiq granitic gneiss thrusting over Um Esh amphibolites, (f) Atalla felsites, (g) granitic gneiss at the southeastern corner of the study area, (h) Um Esh amphibolites and El-Rubbshi serpentinite, (i) Um Had granite, (j) Nubian sandstone, (k) basic metavolcanics, and (l) Massive serpentinite.

rich in mineral deposits. The results can highlight some geochemical aspects of rocks and can aid not only in lithological identification but also in tracing alteration zones and other geochemical characteristics. Furthermore, the significant success of the current approach greatly supports linking of other geophysical datasets (e.g., gravity and magnetic) and evaluate their results for more reliable geological mapping not only in the arid regions but the vegetated ones as well.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Great Thanks to Gamal Abdelhamid for helping in fieldwork. Thanks to ESA for providing Sentinel 2 data. Ali Shebl, Mahmoud Abdellatif and, Mahmoud Ibrahim Abdelaziz are funded by Stipendium Hungaricum scholarship under the joint executive program between Hungary and Egypt. The authors greatly appreciate the referee's valuable comments.

References

- ABD EL MONSEF, M., 2020. Ore Controls and Metallogenesis of Au-Ag Deposits at Atalla Mine, Central Eastern Desert of Egypt. *Acta Geol. Sin. - English Ed.* 94 (5), 1451–1470. <https://doi.org/10.1111/acgs.v94.510.1111/1755-6724.14326>.
- Aero-Service, 1984. Final Operational Report of Airborne Magnetic/Radiation Survey in the Eastern Desert, Egypt. Conducted for the Egyptian General Petroleum Corporation, Aero-Service Division, Houston, Western Geophysical Co., Texas.
- Bachri, I., Hakdaoui, M., Raji, M., Teodoro, A.C., Benbouziane, A., 2019. Machine Learning Algorithms for Automatic Lithological Mapping Using Remote Sensing Data: A Case Study from Souk Arbaa Sahel, Sidi Ifni Inlier, Western Anti-Atlas, Morocco. *ISPRS Int. J. Geo-Information* 8, 248. <https://doi.org/10.3390/ijgi8060248>.
- Charbonneau, B.W., Holman, P.B., Hetu, R.J., 1997. Airborne gamma spectrometer-magnetic-VLF survey of northeastern Alberta. *Bull. Surv. Canada* 107–132.
- Cheng, G., Huang, H., Li, H., Deng, X., Khan, R., SohTamehe, L., Atta, A., Lang, X., Guo, X., 2021. Quantitative Remote Sensing of Metallic Elements for the Qishitan Gold Polymetallic Mining Area, NW China. *Remote Sens.* 2021, Vol. 13, Page 2519 13, 2519. doi: 10.3390/RS13132519.
- Conoco, C., 1987. Geological map of Egypt, scale 1: 500,000-NF 36 NE-Bernice. Egypt. *Egypt. Gen. Pet. Corp. Cairo*.
- Darnley, A.G., 1989. Regional airborne gamma-ray surveys, in: *Proc. Explor. '87*, 3rd Intr. Conf. Geophys. Geochem. Explor. Min. Groundwater. pp. 229–240.
- Deng, Y.F., Song, X.Y., Chen, L.M., Zhou, T., Pirajno, F., Yuan, F., Xie, W., Zhang, D., 2014. Geochemistry of the Huangshandong Ni-Cu deposit in northwestern China: Implications for the formation of magmatic sulfide mineralization in orogenic belts. *Ore Geol. Rev.* 56, 181–198. <https://doi.org/10.1016/J.OREGEOREV.2013.08.012>.
- Dickson, B.L., Scott, K.M., 1997. Interpretation of aerial gamma-ray surveys - adding the geochemical factors.
- Eckstein, B.A., 1989. Evaluation of spline and weighted average interpolation algorithms. *Comput. Geosci.* 15 (1), 79–94. [https://doi.org/10.1016/0098-3004\(89\)90056-3](https://doi.org/10.1016/0098-3004(89)90056-3).
- Elkhateeb, S.O., Abdellatif, M.A.G., 2018. Delineation potential gold mineralization zones in a part of Central Eastern Desert, Egypt using Airborne Magnetic and Radiometric data. *NRIAG J. Astron. Geophys.* 7 (2), 361–376. <https://doi.org/10.1016/j.nrjag.2018.05.010>.
- Grebby, S., Naden, J., Cunningham, D., Tansey, K., 2011. Integrating airborne multispectral imagery and airborne LiDAR data for enhanced lithological mapping in vegetated terrain. *Remote Sens. Environ.* 115 (1), 214–226. <https://doi.org/10.1016/j.rse.2010.08.019>.
- Hamimi, Z., Hagag, W., Kamh, S., El-Araby, A., 2020. Application of remote-sensing techniques in geological and structural mapping of Atalla Shear Zone and Environs, Central Eastern Desert, Egypt. *Arab. J. Geosci.* 13 (11) <https://doi.org/10.1007/s12517-020-05324-8>.
- Harris, J.R., 1989. Clustering of gamma ray spectrometer data using a computer image analysis system. *Stat. Appl. Earth Sci. Geol. Surv. Canada, Canada* 19–31.
- Harris, J.R., Ford, K.L., Charbonneau, B.W., 2014. Application of gamma-ray spectrometer data for lithological mapping in a cordilleran environment. Sekwi Region, NWT. https://doi.org/10.5589/m09-022_35_S12-S30. <https://doi.org/10.5589/M09-022>.
- Harris, J.R., Grunsky, E.C., 2015. Predictive lithological mapping of Canada's North using Random Forest classification applied to geophysical and geochemical data. *Comput. Geosci.* 80, 9–25. <https://doi.org/10.1016/J.CAGEO.2015.03.013>.
- Harris, J.R., Neily, L., Pultz, T., Slaney, V.R., 1987. Principal component analysis of airborne geophysical data for lithologic discrimination using an image analysis system. *Remote Sens. Environ.* 2, 641–657.
- Hunt, G.R., Ashley, R.P., 1979. Spectra of altered rocks in the visible and near infrared. *Econ. Geol.* 74, 1613–1629. <https://doi.org/10.2113/gsecongeo.74.7.1613>.
- Kumar, C., Chatterjee, S., Oommen, T., Guha, A., 2020. Automated lithological mapping by integrating spectral enhancement techniques and machine learning algorithms using AVIRIS-NG hyperspectral data in Gold-bearing granite-greenstone rocks in Hutti, India. *Int. J. Appl. Earth Obs. Geoinf.* 86, 102006. <https://doi.org/10.1016/j.jag.2019.102006>.
- Liesenberg, V., Gloaguen, R., 2012. Evaluating SAR polarization modes at L-band for forest classification purposes in eastern Amazon, Brazil. *Int. J. Appl. Earth Obs. Geoinf.* 21, 122–135. <https://doi.org/10.1016/j.jag.2012.08.016>.
- Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: an applied review. 39 (9), 2784–2817. <https://doi.org/10.1080/01431161.2018.1433343>.
- Mou, D., Wang, Z.-W., Mou, D., Wang, Z.-W., 2015. A comparison of binary and multiclass support vector machine models for volcanic lithology estimation using geophysical log data from Liaohai Basin, China. *Explor. Geophys.* 47, 145–149. <https://doi.org/10.1071/EG14114>.
- Pal, M., Mather, P.M., 2005. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* 26 (5), 1007–1011. <https://doi.org/10.1080/01431160512331314083>.
- Patra, I., Veldi, R.B., 2016. High-resolution Airborne Gamma-ray Spectrometric Data in Geological Mapping-A Case Study from Parts of Shillong Basin, Meghalaya.
- Richards, J.A., Jia, X. (Eds.), 1999. *Remote Sensing Digital Image Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Schetselaar, E.M., Chung, C.J.F., Kim, K.E., 2000. Integration of Landsat TM, Gamma-Ray, Magnetic, and Field Data to Discriminate Lithological Units in Vegetated. *Remote Sens. Environ.* 71, 89–105. [https://doi.org/10.1016/S0034-4257\(99\)00069-3](https://doi.org/10.1016/S0034-4257(99)00069-3).
- Scholkopf, B., Smola, A., 2018. Learning with kernels: support vector machines, regularization, optimization, and beyond.
- Shebl, A., Abdellatif, M., Elkhateeb, S.O., Csámer, Á., 2021. Multisource Data Analysis for Gold Potentiality Mapping of Atalla Area and Its Environs, Central Eastern Desert, Egypt. *Miner.* 2021, Vol. 11, Page 641 11, 641. doi: 10.3390/MIN11060641.
- Shebl, A., Csámer, Á., 2021. Stacked vector multi-source lithologic classification utilizing Machine Learning Algorithms: Data potentiality and dimensionality monitoring. *Remote Sens. Appl. Soc. Environ.* 24, 100643. <https://doi.org/10.1016/j.rsase.2021.100643>.