Article

Not peer-reviewed version

Mapping Dissolved Organic Carbon (DOC) and Organic Iron by Comparing Deep Learning and Linear Regression Techniques Using Sentinel 2 and WorldView 2 Imagery (Byers Peninsula, Marine Antarctica)

Susana del Carmen Fernández *, Rubén Muñiz, Juanjo Peón, Ricardo Rodríguez-Cielos, Jesús Ruíz, Javier F. Calleja

Posted Date: 27 February 2024

doi: 10.20944/preprints202402.1522.v1

Keywords: maritime antarctica; dissolved organic carbon; organic iron; soil mapping; linear regression; deep learning; Sentinel2 images; WordView 2 imagery

Article

Mapping Dissolved Organic Carbon (DOC) and Organic Iron by Comparing Deep Learning and Linear Regression Techniques Using Sentinel 2 and WorldView 2 Imagery (Byers Peninsula, Marine Antarctica)

Susana del Carmen Fernández [1,*], Rubén Muñiz [2], Juanjo Peón [1,3], Ricardo Rodríguez-Cielos [4], Jesús Ruíz [5] and Javier F. Calleja [6]

1 Department of Geology and ICTEA (Instituto Universitario de Ciencias y Tecnologías Aeroespaciales de Asturias). University of Oviedo. Spain; fernandezmsusana@uniovi.es
2 Department of Computer Science. University of Oviedo. Spain; rubenms@uniovi.es
3 juanjopeon@gmail.com
4 Department of Signals, Systems and Radiocommunications (SSR). Polytechnic University of Madrid. Spain; ricardo.rodriguez@upm.es
5 Department of Geography. University of Oviedo. Spain; ruizjesus@uniovi.es
6 Department of Physics. University of Oviedo. Spain; jfcalleja@uniovi.es
* Correspondence: fernandezmsusana@uniovi.es

Abstract: Byers Peninsula is considered one of the largest ice-free areas in Maritime Antarctica. Since 2006, the Spanish Polar Program has taken part in a large number of environmental studies involving effects of climate changes in lives cycles, limnology and microbiology. Soils from maritime Antarctica are generally weakly developed and have chemical, physical and morphological characteristics strongly influenced by the parent material. However, biological activity during the short Antarctic summer promotes intense transference of nutrients and organic matter in areas occupied by different species of birds and marine mammals. To mapping and monitoring those areas with high biological occupation, could be very useful to have models of edaphic properties prepared from satellite images. In this approach, Deep Learning and Linear Regression models of soil properties and spectral indexes as explicative variables were performed. We training models of soil properties closely related to biological activity such as Dissolved Organic Carbon (DOC) and the iron fraction associated with organic matter (Fe). We tested the best approach to model the spatial distribution of DOC, Fe and pH by training models of Linear Regression and Deep Learning over Sentinel2 and WorldView2 images. The most robust models were used to track possible areas with ornithogenic soils as well as areas of the Byers Peninsula that could support the highest biological development.

Keywords: maritime antarctica; dissolved organic carbon; organic iron; soil mapping; linear regression; deep learning; Sentinel2 images; WordView 2 imagery

1. Introduction

Byers Peninsula is located in east end part of Livingston Island in a maritime Antarctic environment, that comprising part of the Antarctic Peninsula and the surrounding islands (Figure 1). Byers Peninsula with about 60 km2 of surface is considered one of the largest ice-free areas in maritime Antarctica and is an Antarctic Special Protected Area (ASPA No. 126), redesignated as Site of Special Scientific Interest (SSSI). The primary reason for the designation of Byers Peninsula as an ASPA and SSSI is to protect the terrestrial and lacustrine habitats within the area. Values protected include the diversity of plant and animal life, many invertebrates, a population of southern elephant seals (Mirounga leonina), small colonies of Antarctic fur seals (Arctocephalus gazella), and a large variety of plants and animals within a relatively small area. During the fourth International Polar Year, Byers

 doi:10.20944/preprints202402.1522.v1

Peninsula was established as an 'International Antarctic Reference Site for Terrestrial, Freshwater and Coastal Ecosystems' . Predict the ecological variations expected after a large scale climate shift was the main goal of research projects as LIMNOPOLAR [1]. During this period, baseline data on terrestrial, limnic and coastal ecosystems were collected, including permafrost characteristics, geomorphology, vegetation extent, limnic diversity and function, marine mammal and bird diversity, microbiology and coastal marine invertebrate diversity [2–13]. Ref. [14] provide a review of all scientific work in the field published between 1957 and 2012.

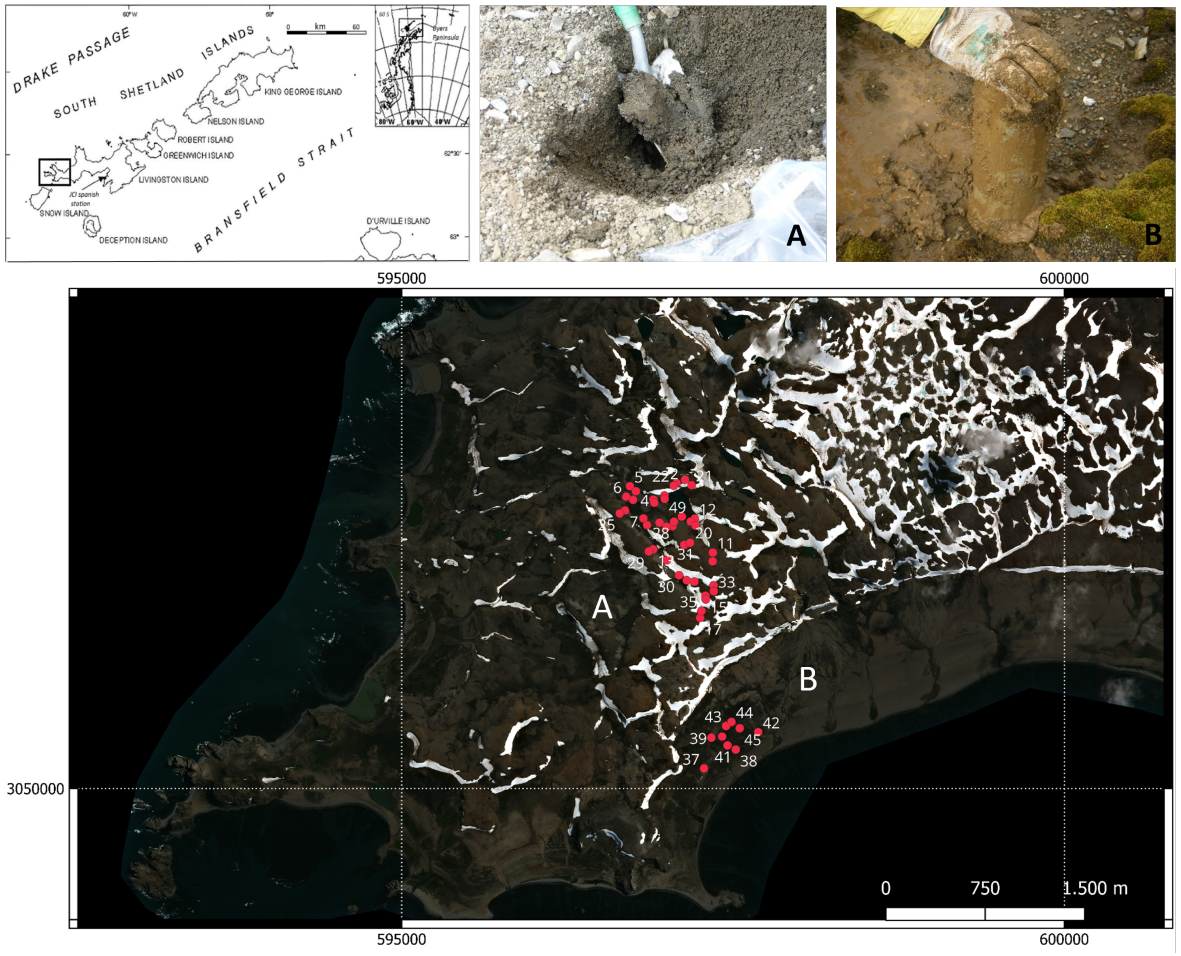

Figure 1. Location of Byers Peninsula. World View 2 image used in this study with location of soil samples and photos showing major soil types sampled. Photo A corresponds to sampling plot A in the image, photo is from beaches (plot B).

Despite the remoteness of this part of Antarctica and the adverse working conditions, information on soil formation in the area is relatively abundant ([15–18] and references therein). The soil survey of the ice-free of Byers Peninsula suggests that mosses are the main cause of the organic soils and were found throughout the ice-free areas of Maritime Antarctica in association with mineral soils [16]. Nevertheless, intense penguin activity on ice-free areas along the Antarctic coast leads to the formation of ornithogenic soils [15,19,20]. According to [19] , guano accumulation in penguin rookeries is the most abundant source of organic matter in the Antarctic terrestrial ecosystem. Skuas (Catharacta sp.) and other flying birds that nest around penguin rookeries also extend the ornithogenic influence farther afield [19].

At upland sites in Byer peninsula with no vegetation or ornithogenic influence, soil chemistry and mineralogy are related mainly to the physical weathering of the substrata [21].However, ornithogenic influence alters soils characteristics leading to soil acidification, leaching of exchangeable bases,

 doi:10.20944/preprints202402.1522.v1

transformation of primary minerals and release of amorphous Fe and Al [18]. The ornithogenic soils are readily distinguishable from the non-ornithogenic soil by lower pH and higher organic matter values [16,22].

Apart from the ornithogenic origin of the organic matter in the Byers soils, the results of the study by [23] shows that the colonization of the substrate by plants is associated with a higher rate of mineral weathering, which has clear repercussions on the mobility and bioavailability of many elements. In this study, higher amounts of the labile forms of organic matter (DOC) and Fe in the soils with vegetation samples than in non-vegetated were found. Although Byers Peninsula is almost free of vegetation, from an Antarctic point of view the vegetation is relatively rich along the more low-lying coastal areas, especially on South Beaches and President Beaches. At these sites, the vegetation is favored by a more advantageous local climate and by nutrient supplies from marine mammals and birds [19]. Deschanylsia antarctica and Colobanthus quitensis, the flowering plants of Antarctica could be found in low-lying areas of the Byers Peninsula, together with mosses, lichens, liverworts, and fungi. Extensive carpets of mosses along President and South Beaches are dominated by Dreparwcladus uncinatus. The progressive colonization of this part of the Byers peninsula by plants may accelerate edaphogenesis, while also substantially affecting the mobility and bioavailability of macro and micronutrients [23]. Sodium pyrophosphate-extracted Fe in Byers soils can be considered as the fraction bound to organic matter (organochelates) and used to estimate the mobility of organo-iron complexes [24] . Thus, in the medium and long term, accelerated mineral weathering by plants and microorganisms may increase the input of bioavailable forms of Fe to coastal and lacustrine waters.

Conducting field surveys in Byers peninsula poses significant difficulties owing to the extreme environmental conditions, in addition to the completed restricted access. Nevertheless, there are numerous works on digital mapping of the Antarctic environment based on satellite remote sensing data [25–29]. The use of images as soil predictors, assuming that remote sensing data is able to give detailed information on the soil properties [30]. This assumption is better applied in areas without continuous vegetation cover, as is the case in Antarctica. [31] used Sentinel2 spectral indexes as predictor variables to performed soil texture maps of all bare soils in Maritime Antarctica.

In this work, we trained models of soil properties closely related to biological activity, such as dissolved organic carbon (DOC) and the iron fraction associated with organic matter (Fe) or chelated iron, and also pH. We tested the best approach to model the spatial distribution of these key soil properties by training linear regression (LRM) and deep learning neural network (DL) models over Sentinel2 and WorldView2 imagery. The most robust models were used to track possible areas of ornithogenic soils and areas of the Byers Peninsula that could support the highest biological development.

The objectives of this research were:

1. To train models of soil properties using optical satellite imagery such as Sentinel2 and WordView2.
2. To search for spectral indices that could be useful for tracking dissolved organic carbon and iron chelates in Byers Peninsula as a training plot for maritime Antarctic periglacial areas.
3. To look for the areas most likely to be biologically colonized. These areas, if accessible, should be the main target of exhaustive inventories and analyses to elucidate the true causes of the increase in their biological activity indicators.

2. Materials and Methods

The Byers Peninsula (Figure 1) is located at the western end of Livingston Island, South Shetland Islands (approx. 62°13’70S, 61°1’60”W), where environmental conditions are more favorable for soil development than in other Antarctic regions [16,23]). This peninsula is a site of special biogeochemical and ecological interest [1]. Biodiversity at this site is higher than at other nearby sites, possibly due to mild environmental conditions, proximity to South America, and potential wind transport of propagules [32].

 doi:10.20944/preprints202402.1522.v1

2.1. Geology and Geomorphology

Glaciers cover most of Livingstone Island, except for Byers Peninsula, which is the largest ice-free area in the South Shetland Island s. The chronology of the deglaciation process has been determined from the dating of deeper lake sediments [33]. The last major deglaciation of the peninsula occurred 5-4 ka BP [34]. The geological substrates of the peninsula are mainly sandstones, shales, microconglomerates, volcanic and volcaniclastic rocks (Upper Jurassic to Lower Cretaceous) with igneous bodies [33,35,36]. Except for the northwestern part of the peninsula, which reaches 268 m.a.s.l., few areas reach elevations above 100 m.a.s.l., except for a few remnant hills (usually volcanic plugs). The highest of these, Chester Cone, forms a prominent feature in the central part of the peninsula, 193 m high. The inland part of the peninsula is a regular undivided platform of about 40 km2 and between 85 and 100 m high, the "Byers Plateau" [37,38], and most of the peninsula is surrounded by extensive beaches.

2.2. Climate, Weathering and Soils

The average annual precipitation in the region can exceed 500 mm and the average annual temperature is about -2 °C. In summer, precipitation is higher than 100 mm and the mean daily temperature is higher than 0.1 °C at sea level [39]. Soils in this part of Antarctica are the result of periglacial and nival processes, as indicated by the patterned soil [15,33]. Weathering in Byers Peninsula soils is controlled by the primary limiting factor in freeze-thaw weathering, the amount of moisture present, rather than temperature [40]. Consequently, the Byers Peninsula soils studied [18,21,23,41] show little development of edaphic and geochemical processes. For example, in soils of the Byers Plateau, [23] found low concentrations of free iron (Fe), generally <1.3%, indicating that mineral weathering is still an incipient process. However, vegetation was found to play an important role in the development of soil-forming processes.

Soil parent materials vary from marine sedimentary to volcanic and volcanoclastic rocks, intruded by igneous bodies. For [2], the soils from northern Byers Peninsula are generally shallow and coarse textured, with low organic matter content. Nevertheless, these authors found ornithogenic soils in the rocky platforms of the northern coastal region. The soils with ornithogenic character have lower pH and higher organic carbon values because the ornithogenic influence leading to soil acidification, leaching of exchangeable bases, transformation of primary minerals and release of amorphous Fe and Al [18].

2.3. Sampling and Analysis

Samples were collected at different locations in two plots (Figure 1; Plots and photos A and B). One plot was located on the Byers Plateau and the other on the Southern Beach. In the Byers Plateau, at an average elevation of 80-100 m.a.s.l., we collected and georeferenced 42 samples. The other 7 samples were collected in the Southern Beaches at sea level. These beaches extend along the south side of Byers Peninsula between Devils Point on the west and Rish Point on the east. In order to obtain comparable data from all sites and for the satellite imagery, the soils were sampled at the surface (0-5 cm). Soils were characterized in the laboratory by analysis of pH, organic carbon extracted with sodium pyrophosphate, and dissolved organic carbon (DOC). In the laboratory, iron associated with the organic fraction was extracted by shaking the samples for 2 h with 0.02M HNO_3 + 30% H_2O_2 at 85°C. Dissolved organic carbon (DOC) was extracted with Milli-Q water (soil:solution ratio, 1:10) for 1 h with continuous shaking at 3°C and 15°C and analyzed in a loop flow analyzer system (Systea).

2.4. Satellite Imagery

In this thesis, Sentinel 2 data were sought for the Byers Peninsula. Several options were analyzed: the use of the Sentinel 2 Global Mosaic (S2GM) and the use of daily imagery. The use of mosaics was discarded as they do not provide an acceptable result due to the high presence of clouds and

 doi:10.20944/preprints202402.1522.v1

snow in the region. Out of more than 200 daily images, in principle only one image could be used, the Sentinel 2-A image from March 28, 2016, at 13:29. The rest have a lot of clouds and snow. The image was corrected for atmosphere using Sen2Cor in SNAP to obtain the L2A level product (BOA reflectivity). For this work, six radiometric indexes are applied to the BOA reflectivity (Table 1). In this sentinel2 image, of the 49 soil samples, 12 fall in areas of the image with clouds or snow, leaving 49-12=37 samples available. Nevertheless, a WorldView-2 image of the Byers Peninsula taken on February 2, 2011 at 13:31 is available for this study. The spatial resolution of the image is 0.5 m for the panchromatic band and 2 m for the others. It is a BOA reflectivity image corrected for atmosphere using ATCOR-2 in ERDAS. We also used this image because it is free of clouds. The 49 soil samples fall in areas without clouds and snow in the WorldView-2 image. The same six radiometric indexes were computed (Table 1).

Table 1. Radiometric indexes used in this study; the expression and Sentinel2 bands involved in their formulation.

Indexes	Expression	Sentinel 2 Bands	WV_2 Bands	Authors
Ferric iron (Fe3)	$\frac{\rho RED}{\rho GREEN}$	B4-VIS - ρRED B3-VIS - ρGREEN	B2-GREEN-ρGREEN B3-RED-ρRED	[42]
Hue	$\arctan \frac{2R-G-B}{30.5} \times (G$ - B)	B2-VIS - ρBLUE B3-VIS - ρGREEN B4-VIS - ρBLUE	B1-BLUE-ρBLUE B2-GREEN-ρGREEN B3-RED-ρRED	[43]
IR550	$IR\_GREEN = \frac{1}{\rho GREEN}$	B3-VIS - ρGREEN	B2-GREEN-ρGREEN	[44]
IR700	$IR\_RED = \frac{1}{\rho RED}$	B5 - ρNIR	B3-RED-ρRED	[44]
Missa Soil Brightness Index (MSBI) v2	$MSBI = 0.406 \times \rho GREEN + 0.600 \times \rho RED + 0.645 \times \rho NIR1 + 0.243 \times \rho NIR2$	B3-VIS - ρGREEN B4-VIS - ρRED B6-NIR - ρNIR1 B8a-NIR - ρNIR2	B0-PAN-ρNIR1 B2-GREEN-ρGREEN B3-RED-ρRED B4-NIR1-ρNIR2	[45]
I/O (Oxides)	IO=$\frac{\rho RED}{\rho BLUE}$	B2-VIS - ρBLUE B4-VIS - ρRED	B1-BLUE-ρBLUE B3-RED-ρRED	[46]

2.5. Modelling

Many authors have tried different techniques to model the spatial distribution of soil properties using explanatory variables. In [25] we can find an extensive review about the works of Digital Soil Mapping published between 2003- 2021. This review is based on 244 works of modelling soil properties at a broad scale (>10,000 km^2). In bare soils of Maritime Antarctic areas the mapping of soil texture by [31] could be a referent for our approach. We had opted by compare two methods, a Deep Learning Neural Network [29] (Figure 2) and Linear regression, to find the best straight line that describes the relationship between the dependent variable and the independent variables. Although it is a straightforward algorithm, linear regression is a powerful tool due to its simplicity and interpretability.

To generate a non-linear regression model we defined and trained a multilayer perceptron [47] (Figure 2). This is a common resource to solving problems where a linear approach is not good enough [48,49]. In summary, the topology of the neural network used in this work is described as follows:

1. Input layer (size = 6)

 doi:10.20944/preprints202402.1522.v1

2. First hidden layer (500 neurons)
3. Second hidden layer (100 neurons)
4. Third hidden layer (50 neurons)
5. Output layer (1 neuron)

The model was trained using the Mean Squared Error (MSE) loss function [50], while the chosen optimizer for this purpose was Adam [51].

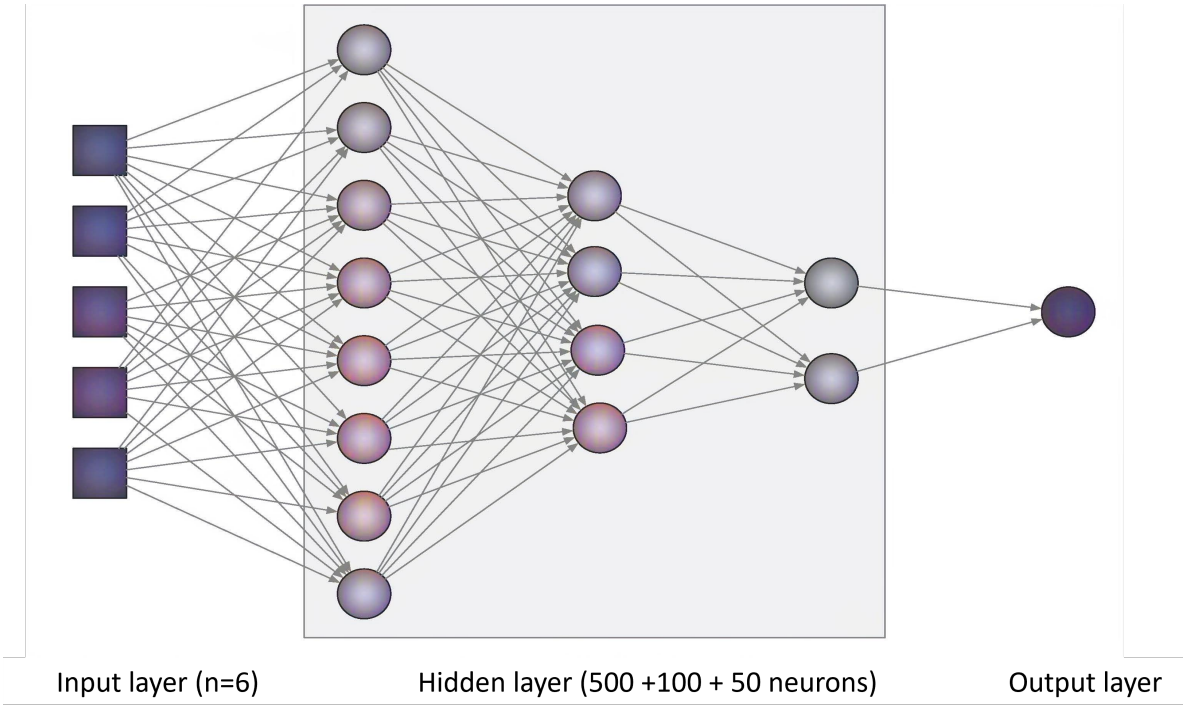

Figure 2. Simplified representation of the neural network topology used.

Multiple linear regression (MLR) was used to establish the relationships between soil properties and explanatory variables. The linear regression model assumes a linear relationship between inputs and outputs, this relationship is represented by a linear equation (Eq. 1):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + e \quad (1)$$

where Y is the dependent variable (soil property), β_0 is the order at the origin or bias, $\beta_1 \ldots \beta_p$ are the coefficients of the independent variables ($X_1 \ldots X_p$ bands and radiometric indexes) that determine the slope of the line; e is the stochastic residual or difference between the model's predictions and the actual values which is assumed to be normally distributed with zero mean and constant variance. The analysis determined the variables using a backward stepwise selection procedure and Akaike information criteria (AIC) to find the model that best explains the data with the fewest parameters.

The training process in linear regression involves adjusting the values of $\beta_0, \beta_1, \beta_2, \ldots \beta_n$ in such a way as to minimize the mean square error (MSE) between predictions and actual values. Once the model has been trained, it can be used to make predictions on new data.

2.6. Validation and Statistical Analysis

The prediction models were validated using leave-one-out cross-validation (LOOCV). The mean absolute error (MAE), root mean square error (RMSE) were used to evaluate the prediction accuracy of different prediction models. The model with the lowest RMSE and MAE values is determined to be the most accurate model. Their equations are expressed as in Equations 2 and 3.

 doi:10.20944/preprints202402.1522.v1

$$MAE = \frac{\sum_{i=0}^{n} |O_i - Pi|^2}{n} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n} (O_i - Pi)^2}{n}} \quad (3)$$

where O_i and P_i are the observed and predicted contents, respectively.

3. Results and Discussion

3.1. Analysis of Soil Properties

In Table 2, we present the results of the laboratory analysis carried out on the 49 soil samples. It is important to note that the soils are poor in organic carbon (average OC 1.14%) with a maximum of 11.92% measured in the sample (40, in the beach; see Figure 1). Other samples with values near to the maximum are samples 38 and 42, taken on the vegetated soils of the beach. But the soils of Byers Plateau (Plot A, Figure 1) are poor in organic carbon. However, even if the organic carbon is very low (minimum 0.31 %) there are always iron chelates (minimum 53.60 mg/kg in sample 26, with a 0.41 of OC %). In addition, when organic carbon is present in soils, about 15% is in the form of dissolved organic carbon. The maximum content of organic carbon is 11.92 g/ 100 g of soil in sample 40 and 3.671g corresponds to DOC. This amount represents about 1/3 of the organic carbon in the sample. This seems to indicate a release of organic carbon and Fe into the Antarctic Spodic cryosol environment, which is described by [16]. All Spodic Cryosols contain large amounts of carbon, nitrogen and iron complexed with organic matter in the topsoil [52]. Mosses and lichens are the organisms that contribute the most C and N (primary source) to soil organic matter on the Byers Peninsula [53]. The relationship between organic carbon and iron complexes in the soils of Byers Peninsula, occupied by mosses and lichens, appears to be very close. The soil with vegetation had the highest concentrations of iron chelates (1768±374.26 mg kg-1). Finally, this work modeled soil pH H2O, which was found to be near neutral in most cases (average pH: 7.3). HHowever, there were significant differences in the surface samples colonized by plants. These samples were significantly more acidic (pH: 5.07±0.1) than the rest of the soils, which had an average pH of 7.32±0.72. [16] found that the soils with ornithogenic origin on Clark Peninsula, near the Wilkes Station (66°15’25"S 110°31’37"E) in Antarctica, have higher pH values than the soils influenced by geological substrata. However, [2] found that ornithogenic soils of the Byers Peninsula have a pH value close to 5, more acid that the soils in Byers Plateau. Ornithogenic soils were not studied in Byers. However, the soils occupied by mosses exhibit the lowest pH values (5.07±0.72). The explanation for this discrepancy with the work of [54] could be related to the nature of the rocks in the substrate. In Byers, pyroclastic and marine sedimentary rocks show the presence of carbonates, so in the case of Byers the substrates are generally neutral with pH values slightly greater than 7.0 [2,21,55].

 doi:10.20944/preprints202402.1522.v1

Table 2. Soil properties measured in the soil samples.

n 49	Mean	Minimum	Maximum	Std.Dev.
DOC (mg/kg)	193.31	0.00	3671.09	656.77
Organic Fe(mg/kg)	286.01	53.60	1768.00	374.26
pH H_2O	7.32	5.07	8.26	0.72
Density (g/cm3)	1.16	0.17	1.50	0.27
OC (%)	1.14	0.31	11.92	2.49
CLAY	15.23	4.13	32.42	5.33
SILT	20.34	4.37	38.59	8.39
SAND	64.45	43.85	87.38	11.26
Mn(g/Kg)	6.40	1.76	10.91	2.04
Ca(mg/kg)	17.17	2.20	29.20	7.62

3.2. Generated Models and Maps of Soil Properties

The Deep Learning model uses 6 explanatory variables (Figure 2) which are the spectral indexes found in Table 1, thus the significance of each variable varied across individual pixels. Therefore, determining the most crucial variables in predicting each soil property using this approach is challenging. On the other hand, although linear regression models have limitations, especially when the relationships between variables are nonlinear, they are a powerful tool due to their simplicity, interpretability, and computational efficiency. Table 3 summarizes the results of the LRM analysis for pH; DOC and Organic Fe.

Of the six spectral indices calculated, only three are included as significant variables in the models. Fe3 and Oxides are explanatory variables in all three models of soil properties, while IR700 (the inverse of Red) is significant only in the model of Dissolved Organic Carbon. The R2 adjustment of these models is moderately low, with an R2 of 0.53 for the pH model, 0.69 for the DOC model, and 0.63 for organic Fe. All variables in the three models passed the significance test (p-level column below 0.05). The beta coefficient value represents the weight of the variable in the model. The Oxides spectral index had the highest values in all three models.

 doi:10.20944/preprints202402.1522.v1

Table 3. Summary of the linear regression models using sentinel2 spectral indices. The values of the standardized coefficients (Beta) and the coefficients (B) of the line of best fit and the significance of the models (p-level) are presented.

	Beta	Std.Err. -of Beta	B	Std.Err. - of B	p-level
Intercept			6.405	0.6747	0.00000
S_Fe3	0.4416	0.1069	2.955	0.7149	0.00015
S_Oxides	-0.7970	0.1069	-1.578	0.2116	0.00000
Variable: H_2O pH; R= 0.74499237; R^2= 0.55501364; Adjusted R^2= 0.53566640; F(2.46)=28.687 p					
Intercept			2897.36	523.5466	0.00000
S_Fe3	-0.7904	0.0960	4796.81	582.4167	0.00000
S_IR700	-0.3992	0.1227	-90.82	27.9104	0.00216
S_Oxides	1.1207	0.1333	2013.34	239.4086	0.00000
Variable: DOC (mg/L); R= 0.84590926; R^2= 0.71556248; Adjusted R^2= 0.69659998; F(3.45)=37.736 p					
Intercept			782.16	312.7250	0.01600
S_Fe3	-0.4729	0.0956	-1638.15	331.3458	0.00001
S_Oxides	0.8587	0.0956	880.5	98.0785	0.00000
Variable: Organic Fe (mg/kg); R= 0.80217191; R^2= 0.64347977; Adjusted R^2= 0.62797889 F(2.46)=41.512 p					

Table 4 presents the linear regression models constructed using spectral indices from the World View2 image. The R2 of fit is lower in the pH and organic Fe models, but higher in the DOC model (R2 0.49 for pH model; R2 0.89 for DOC model and R2 0.41 for organic Fe) than in the models built with the spectral indices on the Sentinel2 image. As with the linear models constructed using the indices calculated in the Sentinel image (refer to Table table:table3), all explanatory variables pass the model significance test with p-levels below 0.05. However, the variables included in the models differ. Only the spectral index (Fe3) is present in the pH and organic Fe models, with different values but the same sign (negative in all six models). The spectral index that measures brightness (MSBI) plays an important role in the DOC and Organic Fe models performed with Worldview2 images.

 doi:10.20944/preprints202402.1522.v1

Table 4. Summary of the linear regression models using sentinel2 spectral indices. The values of the standardized coefficients (Beta) and the coefficients (B) of the line of best fit and the significance of the models (p-level) are presented.

	Beta	Std.Err. -of Beta	B	Std.Err. - of B	p-level
Intercept			2.5146	1.3910	0.0775
WV_Fe3	1.1717	0.2775	5.8471	1.3849	0.0001
WV_IR700	2.5089	0.8228	0.3687	0.1209	0.0039
WV_IR500	-2.0843	0.7100	-0.4143	0.1411	0.0053
WV_MSBI	-0.6200	0.1391	-7.3855	1.6568	0.0001
Variable: H_2O pH; R= 0.73199980; R^2= 0.53582371; Adjusted R^2= 0.49362587; F(2.46)=12.698 p					
Intercept			-728.0	163.55	0.0001
WV_HUE	-0.5351	0.05473	-57338.5	5864.86	0.0000
WV_Oxides	0.1219	0.05341	160.4	70.26	0.0273
WV_MSBI	0.9559	0.05468	10329.5	590.85	0.0000
Variable: DOC (mg/L); R= 0.93861507; R^2= 0.88099825; Adjusted R^2= 0.87306480; F(3.45)=111.05 p					
WV_Fe3	-0.8607	0.2981	-2223.94	770.284	0.0060
WV_IR700	-2.0499	0.8839	-155.98	67.257	0.0251
WV_IR500	1.8132	0.7627	186.60	78.487	0.0218
WV_MSBI	0.7001	0.1494	4318.06	921.569	0.0000
Variable: Organic Fe (mg/kg); R= 0.68146494; R^2= 0.46439447; Adjusted R^2= 0.415703 F(4.44)=9.5375					

3.3. Spatial Distribution of Soil Properties

Since 2011, machine learning has become the most popular way to predict properties in digital soil mapping (DSM). Refs. [56,57] have also seen this trend. For [25] there are some reasons for this increasing use of this type of mapping soils tools. Machine learning and Deep Learning [57] techniques are good at understanding complicated relationships between soil properties and many environmental factors. It usually works better than older statistical methods and mapping techniques. In addition, computers are now more powerful and have new technologies, which makes it easier and faster to create soil maps using large amounts of data. Machine learning is a nonparametric method that does not require any hypothesis on distribution and stationary, which are no longer valid with large spatial extents and legacy data. However, Machine Learning and Deep Learning predictive models focus on prediction performance and overlook the importance of pedological knowledge [25]. In this way, Linear Regression Models as a method that uses structural equations to performed digital soil maps can be use not only for predictive mapping but also for enhancing pedological knowledge [58,59].

The Linear Regressions Models (Tables 3 and 4) of soil properties were extended across the entire Byers Peninsula using the Sentinel2 and Worldview images. In Figures 3–5 we can see the spatial distribution of DOC (mg/L), Organic Fe (mg/kg) and pH models built with deep learning (DL), linear regression on Sentinel2 (LRM_S_) and linear regression on WorldView2 image (LRM_VW). Also, the comparison between the estimated values are presented in a linear graphs. The types of lines have been assigned taking into account the error metric of each model (Table 5). Dark blue color represents the values of the properties in each soil; red line represents the model with lowest MAE and RMSE; dotted gray line represents the models with highest MAE and RMSE. The figures consist of two graphs: one displaying the values for each sample, and the other showing the values sorted from lowest to

 doi:10.20944/preprints202402.1522.v1

highest. The images have not been cropped by the line of coast, but the area of the Sentinel2 image for which the soil property models would be valid has been highlighted.

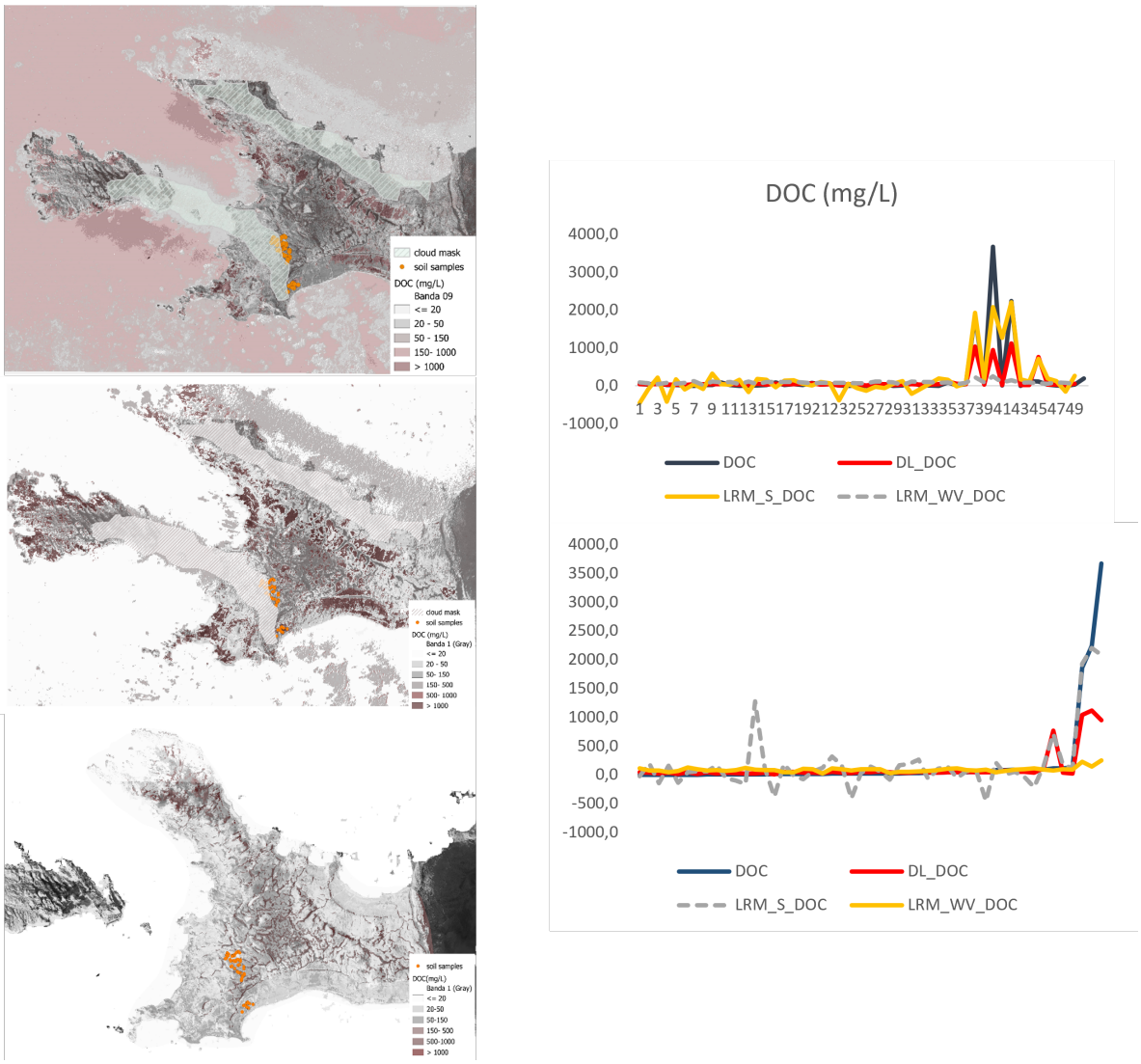

Figure 3. Dissolved organic carbon (DOC mg/L). Models DL_DOC; LRM_S_DOC and LRM_WV_DOC extended to the entire Byers Peninsula: DL and LRM_S were regionalized in Sentinel2 imagery. LRM_WW used WorldView2 image. The smallest and largest values were ordered for each method's measurement and estimation of DOC in the graphs comparison.

 doi:10.20944/preprints202402.1522.v1

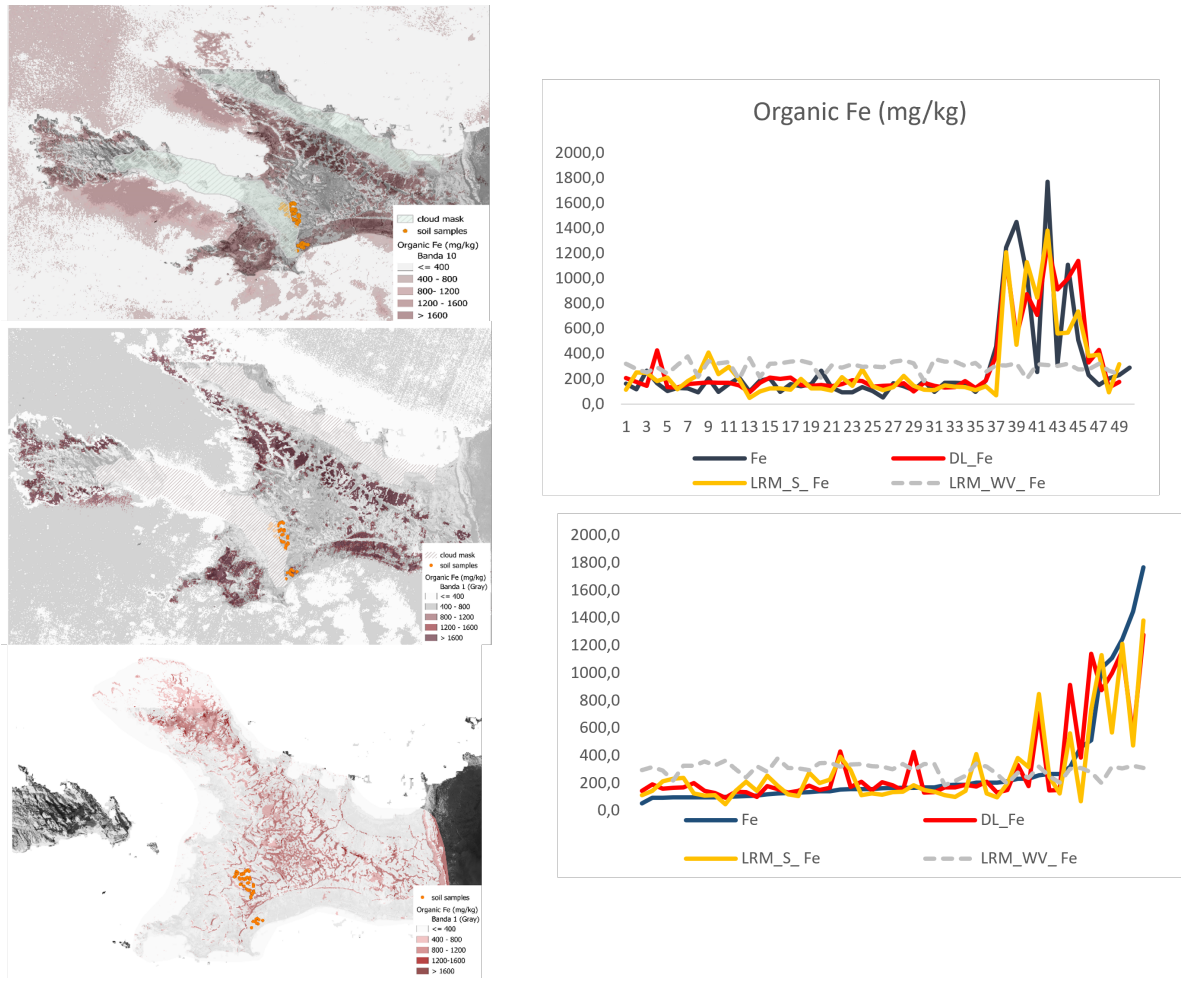

Figure 4. Fe associated with organic matter (mg/kg). Organic Fe (Fe mg/kg). Models DL_Fe; LRM_S_Fe and LRM_WV_Fe extended to the entire Byers Peninsula: DL and LRM_S_ were regionalized in Sentinel2 imagery. LRM_WW used WorldView2 image. The organic Fe was measured and estimated using different methods. The values were then ordered from smallest to largest for each sample and compared in the graphs.

 doi:10.20944/preprints202402.1522.v1

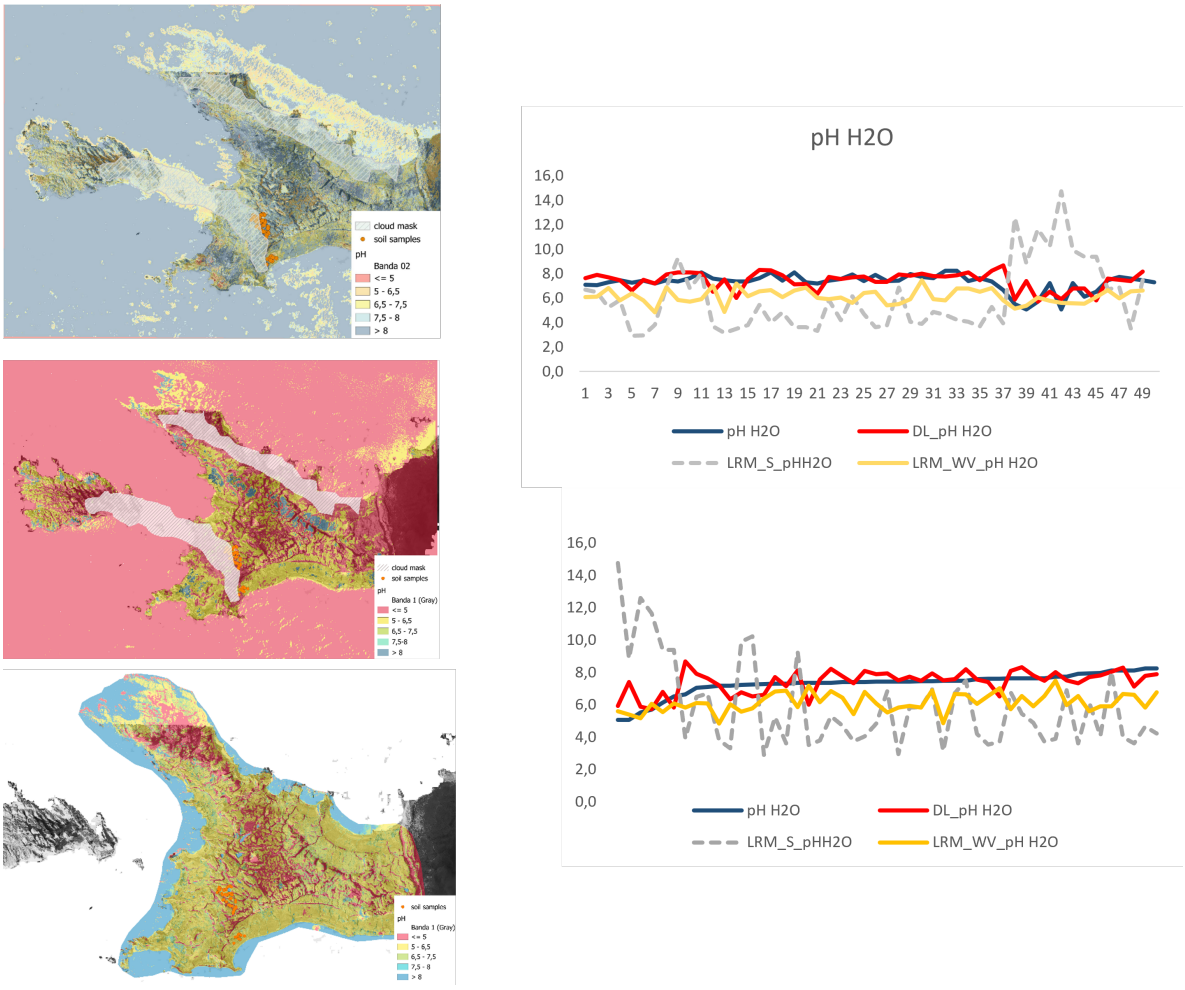

Figure 5. pH models. Models DL_pH; LRM_S_pH and LRM_WV_pH extended to the entire Byers Peninsula: DL and LRM_S_ were regionalized in Sentinel2 imagery. LRM_WW used WorldView2 image. The pH of each sample was measured and estimated using different methods, and the resulting values were sorted from smallest to largest.

Table 5. Error analysis. Deep learning (DL); Linear Regression Models (LRM); WorlsView2 image (WV2); Dissolved Organic Carbon (DOC); Iron associated to organic matter (Fe); R2 Residuals is the correlation coefficient between estimations and residual.

Image	Model	MAE	RMSE	R^2 Residuals
	DL_pH	0.51	0.70	-0.49
Sentinel	LRM_pH	3.04	3.53	-0.99
WV2	LRM_pH	1.21	1.37	-0.43
	DL_DOC	131.87	156.20	0.68
Sentinel	LRM_DOC	189.39	343.23	0.00
WV2	LRM_DOC	202.52	402.12	0.43
	DL_Fe	116.70	209.93	-0.05
Sentinel	LRM_Fe	131.27	219.35	0.00
WV2	LRM_Fe	2689.00	2756.65	-0.80

3.4. Dissolved Organic Carbon (DOC) Models

The dissolved organic carbon has different distribution depending on the model. The DOC models obtained from Sentinel2 images show higher values of DOC in coastal areas and on beaches. This

 doi:10.20944/preprints202402.1522.v1

distribution is consistent with areas that have a higher presence of fauna and vegetation. These results are similar to those obtained in many other studies conducted on the Byers Peninsula [2,16,23,52,60,61]. Mossy soils also exhibit a comparable relationship of significant amounts of organic Fe complexes, indicating that plant activity accelerates the weathering of primary minerals [16,22]. The accompanying graphs for the extended DOC models (Figure 3) show that Deep Learning (red line) best fits the data, while linear regression with the Sentinel2 image performs the worst. However, the linear model with Sentinel2 (gray dotted line) appears to fit the highest DOC values very well. Table 5 displays the MAE and RMSE indices for these DOC models, along with the correlation between residues and soil values. Based on the errors metric, the DL model appears to be the most accurate. However, it is important to note that the residues of this model are highly correlated with the soil values (R2 0.68). The issue with this model is that it does not account for the high DOC soil values. On the other hand, the WorldView model (yellow line) shows no correlation between values and residues.

3.5. Organic Fe Models

In the case of Fe chelates, the highest values are also found on the beaches and along the coast, following the geological patterns described by [2]. These areas are characterized by basalts and pyroclastic rocks. The spatial distribution pattern is similar to that of DOC. The indexes calculated using Sentinel2 images produce the best models, while the estimations using WorldView2 images (gray dotted line) perform poorly (see Table 5).

3.6. pH Models

The spatial distribution of the pH values estimated with DL and LRM calculated in Sentinel2 image shows the influence of the geological substrate. The highest pH values, in dark blue in the maps (Figure 5), are associated with the distribution of basaltic and volcanic rocks [62] and agree with the surface pH values of the profiles of these zones of [2]. The pH values below 5.0 are also consistent with the location of the ornithogenic soils described by this author in Byers. On the other hand, the model created using the WorldView2 image displays a distribution of pH values around 5.0, which are present in all the low-lying areas associated with the valley bottom deposits [23,33]. However, only the soil samples taken from Plot A (38 to 45) show pH values below 6.0. According to the studies on Byers soils consulted in this work, the soils in these depressed areas, covered with fine sediments and incipient vegetation, present the lowest pH values. However, these results were not found in the soils of Byers Plateau (Plot A, Figure 1), as they show pH values over 6.5. However, the pH model created using the indexes of the Worldview2 image, despite producing the worst linear fits (R2 0.49), provides spatial estimations that agree with the pH values measured in various studies of Byers' soils (e.g.:[18,21,41]. The graphs in the figure indicate that this model tends to underestimate pH values in soils of the Byers Plateau (up to sample 37, as shown in Figure 1. However, it accurately adjusts to lower pH values associated with beach areas (samples between 37 and 49, as shown in Figure 1) where there is more vegetation development ([23]). The model with the worst pH estimation is LMR_S_pH, with an MAE of 3.04 and RMSE of 3.53. The residuals are highly correlated with the estimations, with an R2 of 0.99. This model overestimates low pH values and underestimates higher values, making it inaccurate in its distribution of values. However, it does present a better linear fit (R2 0.54) than the pH model built with spectral index in the WorldView image (R2 0.49). On the other hand, LRM_WV_pH presents lower MAE and RMSE values and the lowest residual correlation (-0.43, see Table 5).

We utilized spectral indexes from Sentinel2 and worldView2 images as soil predictors. Remote sensing data can provide detailed information on soil properties, as bare surface reflectance is an intrinsic property of the soil material and may indicate soil attributes such as mineralogy [30]. This assumption is more applicable in regions with limited vegetation, such as Antarctica. [31] utilized Sentinel2 spectral indexes, such as the Soil Adjusted Vegetation Index (SAVI) and Normalized Difference Vegetation Index (NDVI), as predictor variables to map soil texture models in the bare soils of Maritime Antarctica. A Boolean (0-1) vegetation layer was obtained using these spectral indexes,

 doi:10.20944/preprints202402.1522.v1

and the CHELSA net primary productivity (NPP) was used to represent the influence of vegetation or fauna activity, mainly penguins. However, in Byers Peninsula the vegetation cover is very scarce or practically non-existent [12], so we considered other radiometric indices more suitable to carry out this analysis. The radiometric indices utilized in this analysis are closely linked to the geochemical composition and color of the bare soils on the Byers Peninsula (see Table 1).

3.7. Searching Areas of Biological Occupation

This section provides examples of the study's implementation. Figure 6 displays a map of the Byers Peninsula, highlighting areas with concentrations of DOC and organic iron within the range of values found in the studied vegetated soils (soils 38 to 45; see Figure 1) using geochemical criteria. To make this classification, the best DOC model (DL) and the best organic Fe (LRM_S_Fe) model have been used (Table 5). Areas that meet the geochemical requirements specified in the map legend may be preferred for designing wildlife or vegetation studies in Byers. The second example could be applied in the search for ornithogenic soils that have not yet been discovered or studied. In this new classification, the models for pH, DOC, and organic Fe that have provided the lowest values of MAE and RMSE were used (Table 5). The values of the variables have been established following the analyses of ornithogenics soil profiles studied by [2] in Byers peninsula.

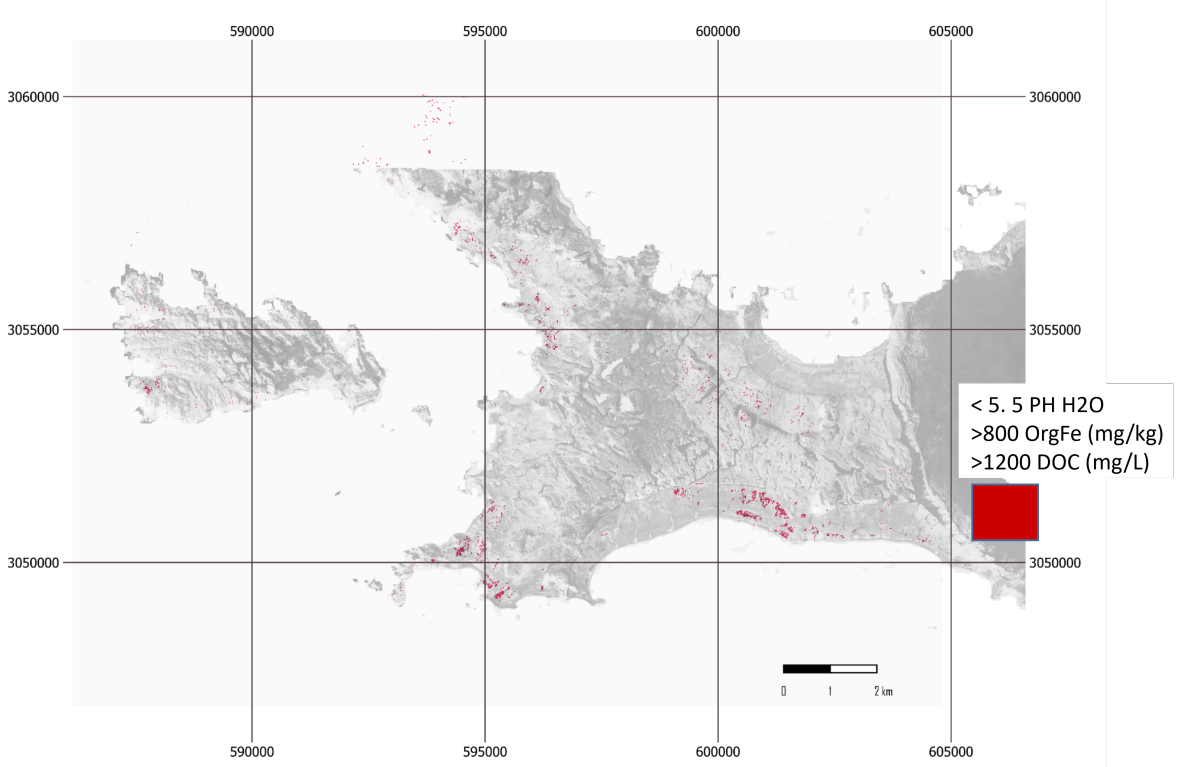

Figure 6. Map of the areas more prone to be colonise by vegetation or occupies by fauna.

 doi:10.20944/preprints202402.1522.v1

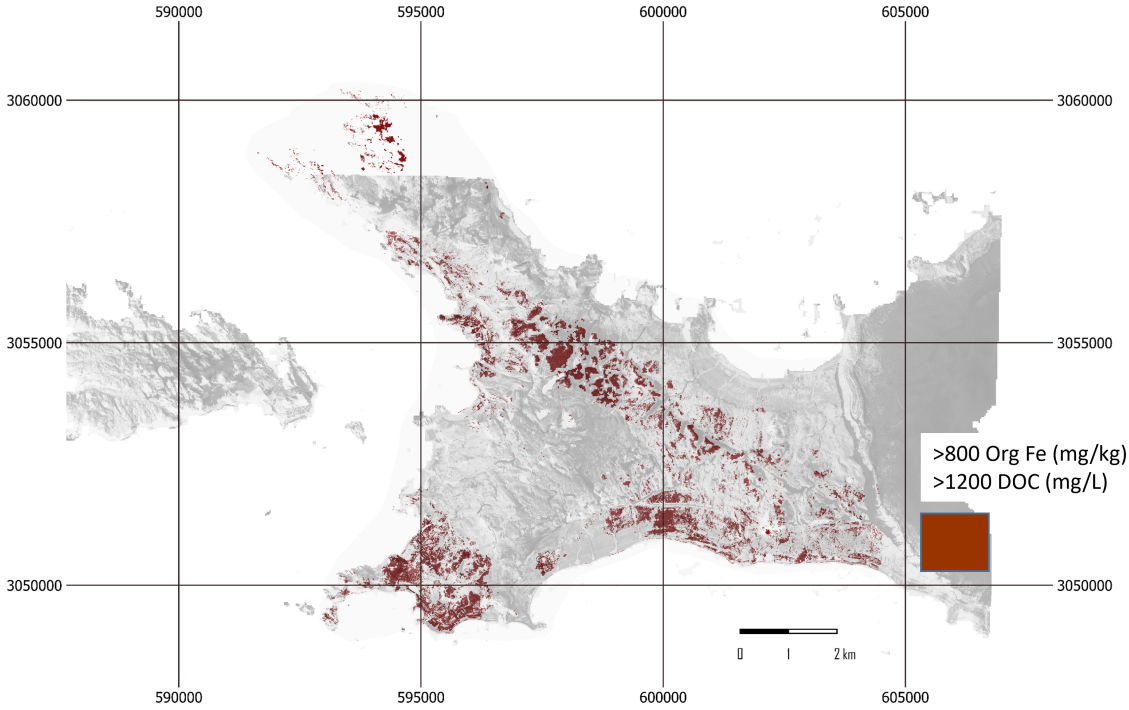

Figure 7. Map of the areas with geochemical evidences of ornithogenic soils, based on [2].

4. Conclusions

This study focused on Byers Peninsula, a significant ice-free region in Maritime Antarctica. The Spanish Polar Program has been actively involved in environmental research in this area since 2006. The study emphasizes the importance of using a non-invasive method to locate and monitor areas of Byers Peninsula that experience significant nutrient and organic matter transfer due to biological activity during the brief Antarctic summer. Various bird and marine mammal species are primarily responsible for this activity. The research aims to determine the most effective method for digital mapping (DSM) using optical satellite images. Soil properties such as dissolved organic carbon, organic-complexes of iron, and pH, which are linked to biological activity, will also be analyzed. This study aimed to identify robust models for tracking areas with ornithogenic soils and regions in Byers Peninsula capable of supporting substantial biological development by comparing the performance of Linear Regression and Deep Learning models. The Deep Learning models (DL) provided the best fits of the modelled properties, showing the lowest MAE and RMSE. However, the models exhibited significant correlations between the values of the variables and the residuals. The analysis suggests that DL models perform well in predicting values near the mean of the modeled variables, with the prediction error increasing or decreasing as the variable value deviates from the mean. In contrast, LRM models do not exhibit this issue. Additionally, the study found that using spectral indexes with Sentinel2 images yielded significantly better results. Perhaps the best contribution of this study is that Sentinel2 images, which have a high temporal resolution and are also free, provide better estimates regardless of the DSM method used.

Author Contributions: Conceptualization, S.C. and R.R.; methodology, S.C. and R.R.; software, R.M.; validation, S.C. and J.R.; formal analysis, S.C, R.M. and J.P.; investigation, S.C., J.R and J.P.; resources, J.R. and R.M.; data curation, S.C. and R.M.; writing—original draft preparation, S.C.; writing—review and editing, R.M.; visualization, S.C., RM. and R.R.; supervision, S.C. and R.M.; project administration, S.C and JF.C.; funding acquisition, S.C. and JF.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by Ministerio de Ciencia e Innovación: PID2021-127060OB-I00, PID2020-113051RB-C31, CTM2017-84441-R, and CTM2014-52021-R.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study, as stated in the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Quesada, A.; Camacho, A.; Rochera, C.; Velázquez, D. Byers Peninsula: a reference site for coastal, terrestrial and limnetic ecosystem studies in maritime Antarctica. *Polar Science* **2009**, *3*, 181–187.
2. Moura, P.A.; Francelino, M.R.; Schaefer, C.E.G.; Simas, F.N.; de Mendonça, B.A. Distribution and characterization of soils and landform relationships in Byers Peninsula, Livingston Island, Maritime Antarctica. *Geomorphology* **2012**, *155-156*, 45–54. Advances in Permafrost and Periglacial research in Antarctica, https://doi.org/https://doi.org/10.1016/j.geomorph.2011.12.011.
3. López-Bueno, A.; Rastrojo, A.; Peiró, R.; Arenas, M.; Alcamí, A. Ecological connectivity shapes quasispecies structure of RNA viruses in an Antarctic lake. *Mol Ecol* **2015**, *24*, 4812–4825.
4. Barbosa, A.; De Mas, E.; Benzal, J.; Diaz, J.I.; Motas, M.; Jerez, S.; Pertierra, L.; Benayas, J.; Justel, A.; Lauzurica, P.; et al. Pollution and physiological variability in gentoo penguins at two rookeries with different levels of human visitation. *Antarctic Science* **2013**, *25*, 329–338.
5. de Pablo, M.A.; Ramos, M.; Molina, A. Thermal characterization of the active layer at the Limnopolar Lake CALM-S site on Byers Peninsula (Livingston Island), Antarctica. *Solid Earth* **2014**, *5*, 721–739. https://doi.org/10.5194/se-5-721-2014.
6. Emslie, S.D.; Polito, M.J.; Patterson, W.P. Stable isotope analysis of ancient and modern gentoo penguin egg membrane and the krill surplus hypothesis in Antarctica. *Antarctic Science* **2013**, *25*, 213–218.
7. Kopalová, K.; Van de Vijver, B. Structure and ecology of freshwater benthic diatom communities from Byers Peninsula, Livingston Island, South Shetland Islands. *Antarctic science* **2013**, *25*, 239–253.
8. Lyons, W.B.; Welch, K.; Welch, S.; Camacho, A.; Rochera, C.; Michaud, L.; Dewit, R.; Carey, A. Geochemistry of streams from byers peninsula, Livingston Island. *Antarctic Science* **2013**, *25*, 181–190.
9. Nakai, R.; Shibuya, E.; Justel, A.; Rico, E.; Quesada, A.; Kobayashi, F.; Iwasaka, Y.; Shi, G.Y.; Amano, Y.; Iwatsuki, T.; et al. Phylogeographic analysis of filterable bacteria with special reference to Rhizobiales strains that occur in cryospheric habitats. *Antarctic Science* **2013**, *25*, 219–228.
10. Toro, M.; Granados, I.; Pla, S.; Giralt, S.; Antoniades, D.; Galán, L.; Cortizas, A.M.; Lim, H.S.; Appleby, P.G. Chronostratigraphy of the sedimentary record of limnopolar lake, Byers peninsula, Livingston island, Antarctica. *Antarctic Science* **2013**, *25*, 198–212.
11. Velázquez, D.; Lezcano, M.Á.; Frias, A.; Quesada, A. Ecological relationships and stoichiometry within a Maritime Antarctic watershed. *Antarctic Science* **2013**, *25*, 191–197.
12. Vera, M.L.; Fernández-Teruel, T.; Quesada, A. Distribution and reproductive capacity of Deschampsia antarctica and Colobanthus quitensis on Byers Peninsula, Livingston Island, South Shetland Islands, Antarctica. *Antarctic Science* **2013**, *25*, 292–302.
13. Villaescusa, J.A.; Casamayor, E.O.; Rochera, C.; Quesada, A.; Michaud, L.; Camacho, A. Heterogeneous vertical structure of the bacterioplankton community in a non-stratified Antarctic lake. *Antarctic Science* **2013**, *25*, 229–238.
14. Benayas, J.; Pertierra, L.R.; Tejedo, P.; Lara, F.; Bermudez, O.; Hughes, K.A.; Quesada, A. A review of scientific research trends within ASPA No. 126 Byers Peninsula, South Shetland Islands, Antarctica. *Antarctic Science* **2013**, *25*, 128 – 145.
15. Campbell, I.; Claridge, G. *Antarctica: Soils, Weathering Processes and Environment*; ISSN, Elsevier Science, 1987.
16. Beyer, L. Properties, formation, and geo-ecological significance of organic soils in the coastal region of East Antarctica (Wilkes Land). *CATENA* **2000**, *39*, 79–93. https://doi.org/https://doi.org/10.1016/S0341-8162(99)00090-9.
17. Ugolini, F.; Bockheim, J. Antarctic soils and soil formation in a changing environment: A review. *Geoderma* **2008**, *144*, 1–8. Antarctic Soils and Soil Forming Processes in a Changing Environment, https://doi.org/https://doi.org/10.1016/j.geoderma.2007.10.005.
18. Simas, F.N.B.; Schaefer, C.E.G.R.; Michel, R.F.; Francelino, M.R.; Bockheim, J.G., Soils of the South Orkney and South Shetland Islands, Antarctica. In *The Soils of Antarctica*; Springer International Publishing: Cham, 2015; pp. 227–273. https://doi.org/10.1007/978-3-319-05497-1_13.
19. Ugolini, F. Ornithogenic soils of Antarctica. *Antarctic terrestrial biology* **1972**, *20*, 181–193.

20. Tatur, A.; Myrcha, A. Ornithogenic ecosystems in the maritime Antarctic-formation, development and disintegration. *Ecological Studies* **2002**, pp. 161–186.
21. Navas, A.; López-Martínez, J.; Casas, J.; Machín, J.; Durán, J.J.; Serrano, E.; Cuchi, J.A. Soil characteristics along a transect on raised marine surfaces on Byers Peninsula, Livingston Island, South Shetland Islands. *Antarctica: contributions to global earth sciences* **2006**, pp. 467–473.
22. Bölter, M.; Kandeler, E. Microorganisms and microbial processes in Antarctic soils. In *Cryosols: Permafrost-Affected Soils*; Springer, 2004; pp. 557–572.
23. Otero, X.; Fernández, S.; de Pablo Hernandez, M.; Nizoli, E.; Quesada, A. Plant communities as a key factor in biogeochemical processes involving micronutrients (Fe, Mn, Co, and Cu) in Antarctic soils (Byers Peninsula, maritime Antarctica). *Geoderma* **2013**, *195*, 145–154.
24. Gamble, A.; Howe, J.; Delaney, D.; Van Santen, E.; Yates, R. Iron chelates alleviate iron chlorosis in soybean on high pH soils. *Agronomy Journal* **2014**, *106*, 1251–1257.
25. Chen, S.; Arrouays, D.; Leatitia Mulder, V.; Poggio, L.; Minasny, B.; Roudier, P.; Libohova, Z.; Lagacherie, P.; Shi, Z.; Hannam, J.; et al. Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. *Geoderma* **2022**, *409*, 115567. https://doi.org/https://doi.org/10.1016/j.geoderma.2021.115567.
26. Turner, D.; Lucieer, A.; Malenovský, Z.; King, D.; Robinson, S.A. Assessment of Antarctic moss health from multi-sensor UAS imagery with Random Forest Modelling. *International Journal of Applied Earth Observation and Geoinformation* **2018**, *68*, 168–179. https://doi.org/https://doi.org/10.1016/j.jag.2018.01.004.
27. Román, A.; Tovar-Sánchez, A.; Fernández-Marín, B.; Navarro, G.; Barbero, L. Characterization of an antarctic penguin colony ecosystem using high-resolution UAV hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation* **2023**, *125*, 103565. https://doi.org/https://doi.org/10.1016/j.jag.2023.103565.
28. Zmarz, A.; Rodzewicz, M.; Dąbski, M.; Karsznia, I.; Korczak-Abshire, M.; Chwedorzewska, K.J. Application of UAV BVLOS remote sensing data for multi-faceted analysis of Antarctic ecosystem. *Remote Sensing of Environment* **2018**, *217*, 375–388. https://doi.org/https://doi.org/10.1016/j.rse.2018.08.031.
29. Fernández, S.; Muñiz, R.; Peón, J.; Rodríguez-Cielos, R.; Pisabarro, A.; Calleja, J. Machine learning and linear regression models for mapping soil properties and albedo in periglacial areas using Sentinel imagery (Byers Peninsula, Marine Antarctica). *SENSORS* **2024**. In press.
30. Minasny, B.; McBratney, A. Digital soil mapping: A brief history and some lessons. *Geoderma* **2016**, *264*, 301–311. Soil mapping, classification, and modelling: history and future directions, https://doi.org/https://doi.org/10.1016/j.geoderma.2015.07.017.
31. Siqueira, R.G.; Moquedace, C.M.; Francelino, M.R.; Schaefer, C.E.; Fernandes-Filho, E.I. Machine learning applied for Antarctic soil mapping: Spatial prediction of soil texture for Maritime Antarctica and Northern Antarctic Peninsula. *Geoderma* **2023**, *432*, 116405. https://doi.org/https://doi.org/10.1016/j.geoderma.2023.116405.
32. Quayle, W.C.; Convey, P.; Peck, L.S.; Ellis-Evans, C.J.; Butler, H.G.; Peat, H.J. Ecological responses of maritime Antarctic lakes to regional climate change. *Antarctic Research Series* **2003**, *79*, 159–170.
33. Lopez-Martinez, J.; Thomson, M.; Arche, A.; Bjorck, S.; Ellis-Evans, J.C.; Hatway, B.; Hernández-Cifuentes, F.; Hjort, C.; Ingolfsson, O.; Ising, J.; et al. Geomorphological map of Byers peninsula, livingston island. *British Antarctic Survey* **1996**.
34. Björck, S.; Hjort, C.; Ingólfsson, O.; Zale, R.; Ising, J. Holocene deglaciation chronology from lake sediments. *Geomorphological map of Byers Peninsula, Livingston Island. BAS GEOMAP Series, Sheet* **1996**, pp. 49–51.
35. Hobbs, G. The geology of the South Shetland Islands: IV. The geology of Livingston Island. *British Antarctic Survey* **1968**.
36. Hathway, B. Nonmarine sedimentation in an Early Cretaceous extensional continental-margin arc, Byers Peninsula, Livingston Island, South Shetland Islands. *Journal of Sedimentary Research* **1997**, *67*, 686–697.
37. van Zinderen Bakker, E.M. *Palaeoecology of Africa & of the Surrounding Islands & Antarctica*; Vol. 16, AA Balkema, 1966.
38. Simms, A.R.; Milliken, K.T.; Anderson, J.B.; Wellner, J.S. The marine record of deglaciation of the South Shetland Islands, Antarctica since the Last Glacial Maximum. *Quaternary Science Reviews* **2011**, *30*, 1583–1601. https://doi.org/https://doi.org/10.1016/j.quascirev.2011.03.018.

 doi:10.20944/preprints202402.1522.v1

39. Rakusa-Suszczewski, S., King George Island — South Shetland Islands, Maritime Antarctic. In *Geoecology of Antarctic Ice-Free Coastal Landscapes*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2002; pp. 23–39. https://doi.org/10.1007/978-3-642-56318-8_3.
40. Hall, M.; Allinson, D. Assessing the effects of soil grading on the moisture content-dependent thermal conductivity of stabilised rammed earth materials. *Applied Thermal Engineering* **2009**, *29*, 740–747. https://doi.org/https://doi.org/10.1016/j.applthermaleng.2008.03.051.
41. Navas, A.; López-Martínez, J.; Casas, J.; Machín, J.; Durán, J.J.; Serrano, E.; Cuchi, J.A.; Mink, S. Soil characteristics on varying lithological substrates in the South Shetland Islands, maritime Antarctica. *Geoderma* **2008**, *144*, 123–139.
42. Mars, J.C.; Rowan, L.C. ASTER spectral analysis and lithologic mapping of the Khanneshin carbonatite volcano, Afghanistan. *Geosphere* **2011**, *7*, 276–289, [https://pubs.geoscienceworld.org/gsa/geosphere/article-pdf/7/1/276/3341201/276.pdf]. https://doi.org/10.1130/GES00630.1.
43. Escadafal, R.; Girard, M.C.; Courault, D. Munsell soil color and soil reflectance in the visible spectral bands of landsat MSS and TM data. *Remote Sensing of Environment* **1989**, *27*, 37–46. https://doi.org/https://doi.org/10.1016/0034-4257(89)90035-7.
44. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment* **1996**, *58*, 289–298. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00072-7.
45. Misra, A.; Prasad, R.C.; Chauhan, V.S.; Srilakshmi, B. A theoretical model for the electromagnetic radiation emission during plastic deformation and crack propagation in metallic materials. *International journal of fracture* **2007**, *145*, 99–121.
46. Hewson, R.; Robson, D.; Mauger, A.; Cudahy, T.; Thomas, M.; Jones, S. Using the Geoscience Australia-CSIRO ASTER maps and airborne geophysics to explore Australian geoscience. *Journal of Spatial Science* **2015**, *60*, 207–231.
47. Murtagh, F. Multilayer perceptrons for classification and regression. *Neurocomputing* **1991**, *2*, 183–197.
48. Muñiz, R.; Cuevas-Valdés, M.; de la Roza-Delgado, B. Milk quality control requirement evaluation using a handheld near infrared reflectance spectrophotometer and a bespoke mobile application. *Journal of Food Composition and Analysis* **2020**, *86*, 103388.
49. Gaudart, J.; Giusiano, B.; Huiart, L. Comparison of the performance of multi-layer perceptron and linear regression for epidemiological data. *Computational statistics & data analysis* **2004**, *44*, 547–570.
50. Hodson, T.O.; Over, T.M.; Foks, S.S. Mean squared error, deconstructed. *Journal of Advances in Modeling Earth Systems* **2021**, *13*, e2021MS002681.
51. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
52. Beyer, L.; Bölter, M. Chemical and biological properties, formation, occurrence and classification of Spodic Cryosols in a terrestrial ecosystem of East Antarctica (Wilkes Land). *CATENA* **2000**, *39*, 95–119. https://doi.org/https://doi.org/10.1016/S0341-8162(99)00089-2.
53. da Silva, J.P.; Lelis Leal de Souza, J.J.; Mercês Barros Soares, E.; Schaefer, C.E.G. Soil organic matter accumulation before, during, and after the last glacial maximum in Byers Peninsula, Maritime Antarctica. *Geoderma* **2022**, *428*, 116221. https://doi.org/https://doi.org/10.1016/j.geoderma.2022.116221.
54. Blume, H.P.; Kuhn, D.; Bölter, M. Soils and soilscapes. In *Geoecology of Antarctic ice-free coastal landscapes*; Springer, 2002; pp. 91–113.
55. Otero, X.L.; Fernandez, S.; Hernández, M.A.; Nizoli, É.C.; Quesada, A. Plant communities as a key factor in biogeochemical processes involving micronutrients (Fe, Mn, Co, and Cu) in Antarctic soils (Byers Peninsula, maritime Antarctica). *Geoderma* **2013**, *195*, 145–154.
56. Arrouays, D.; Poggio, L.; Salazar Guerrero, O.A.; Mulder, V.L. Digital soil mapping and GlobalSoilMap. Main advances and ways forward. *Geoderma Regional* **2020**, *21*, e00265. https://doi.org/https://doi.org/10.1016/j.geodrs.2020.e00265.
57. Padarian, J.; Minasny, B.; McBratney, A. Using deep learning to predict soil properties from regional spectral data. *Geoderma Regional* **2019**, *16*, e00198. https://doi.org/https://doi.org/10.1016/j.geodrs.2018.e00198.
58. Wadoux, A.M.C.; Minasny, B.; McBratney, A.B. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews* **2020**, *210*, 103359. https://doi.org/https://doi.org/10.1016/j.earscirev.2020.103359.

 doi:10.20944/preprints202402.1522.v1

59. Ma, Y.; Minasny, B.; McBratney, A.; Poggio, L.; Fajardo, M. Predicting soil properties in 3D: Should depth be a covariate? *Geoderma* **2021**, *383*, 114794. https://doi.org/https://doi.org/10.1016/j.geoderma.2020.114794.
60. Abakumov, E.; Gagarina, E.; Sapega, V.; Vlasov, D.Y. Micromorphological features of the fine earth and skeletal fractions of soils of West Antarctica in the areas of Russian Antarctic stations. *Eurasian Soil Science* **2013**, *46*, 1219–1229.
61. Silvero, N.E.; Demattê, J.A.; de Souza Vieira, J.; de Oliveira Mello, F.A.; Amorim, M.T.A.; Poppiel, R.R.; de Sousa Mendes, W.; Bonfatti, B.R. Soil property maps with satellite images at multiple scales and its impact on management and classification. *Geoderma* **2021**, *397*, 115089. https://doi.org/https://doi.org/10.1016/j.geoderma.2021.115089.
62. Hathway, B.; Lomas, S. The jurassic–lower cretaceous Byers group, South Shetland islands, Antarctica: Revised stratigraphy and regional correlations. *Cretaceous Research* **1998**, *19*, 43–67.