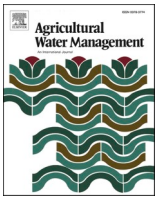
Agricultural Water Management 289 (2023) 108524

Contents lists available at ScienceDirect

Agricultural Water Management

journal homepage: www.elsevier.com/locate/agwat



Sentinel-2-based predictions of soil depth to inform water and nutrient retention strategies in dryland wheat

Simon I. Fordyce [a,*,1], Patrick M. Carr [a,2,3], Clain Jones [b,4,5], Jed O. Eberly [a,2,6], W. Adam Sigler [b,4,7], Stephanie Ewing [b,4,8], Scott L. Powell [b,4,9]

[a] Central Agricultural Research Center, Department of Research Centers, Montana State University, USA
[b] Department of Land Resources and Environment Sciences, Montana State University, USA

ARTICLE INFO

Handling Editor - B.E. Clothier

Keywords:
Machine learning
NDVI
Nitrate leaching
Precision agriculture
Soil thickness
Soil water storage capacity

ABSTRACT

The thickness or depth of fine-textured soil (z_f) dominates water storage capacity and exerts a control on nutrient leaching in semi-arid agroecosystems. At small pixel sizes (< 1 m; 'fine resolution'), the normalized difference vegetation index (NDVI) of cereal crops during senescence (Zadoks Growth Stages [ZGS] 90–93) offers a promising alternative to destructive sampling of z_f using soil pits. However, it is unclear whether correlations between z_f and NDVI exist (a) at larger pixel sizes (1–10 m; 'intermediate resolution') and (b) across field boundaries. The relationship of z_f to NDVI of wheat (*Triticum aestivum* L.) was tested using images from a combination of multispectral sensors and fields in central Montana. NDVI was derived for one field using sensors of fine and intermediate spatial resolution and for three fields using intermediate resolution sensors only. Among images acquired during crop senescence, z_f was correlated with NDVI ($p < 0.05$) independent of sensor ($p = 0.22$) and field ($p = 0.94$). The z_f relationship to NDVI was highly dependent on acquisition day ($p < 0.05$), but only when pre-senescence (ZGS $\leq$ 89) images were included in the analysis. Results indicate that cereal crop NDVI of intermediate resolution can be used to characterize z_f across field boundaries if image acquisition occurs during crop senescence. Based on these findings, an empirical index was derived from multi-temporal Sentinel-2 imagery to estimate z_f on fields in and beyond the study area.

1. Introduction

The ability of soil to store water is a function of texture, rock content, and depth, resulting in a storage capacity that is critical for crop production, particularly in non-irrigated, semi-arid systems. The thickness or depth of fine-textured material (z_f) in the upper profile retains water against deeper percolation and dominates soil-water storage capacity (Sigler et al., 2020). In this sense, z_f impacts within-field heterogeneity of key agronomic variables by virtue of water storage. For example, z_f affects wheat grain yield (Grylls et al., 1997; Bestwick, 2016; John et al., 2017; Krüger et al., 2020) and grain protein concentration (Bestwick, 2016; John et al., 2017), as well as related edaphic properties including precipitation storage efficiency of fallow (Vomocil and Ramig, 1984; Sigler et al., 2020; Carr et al., 2021), inorganic nitrogen (N) changes during the off-season (Jones et al., 2011; Fordyce, 2022), nitrate leaching (John et al., 2017; Sigler et al., 2018; Sigler et al., 2020), and denitrification (Sigler et al., 2022).

In some environments, z_f is highly variable over short distances

* Correspondence to: 52583 US Highway 87, Moccasin, MT 59462, USA.
E-mail address: simon.fordyce@montana.edu (S.I. Fordyce).
[1] ORCID: 0000–0002-2615–349X
[2] 52583 US Highway 87, Moccasin, MT 59462
[3] ORCID: 0000–0001-9243–2600
[4] Leon Johnson Hall, Montana State University, Bozeman, MT 59717–3120
[5] ORCID: 0000–0003-3688–6256
[6] ORCID: 0000–0002-3963–5367
[7] ORCID: 0000–0002-4815–0580
[8] ORCID: 0000–0003-0713–4266
[9] ORCID: 0000–0002-5752–0408

https://doi.org/10.1016/j.agwat.2023.108524
Received 26 April 2023; Received in revised form 18 September 2023; Accepted 20 September 2023
Available online 22 September 2023
0378-3774/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

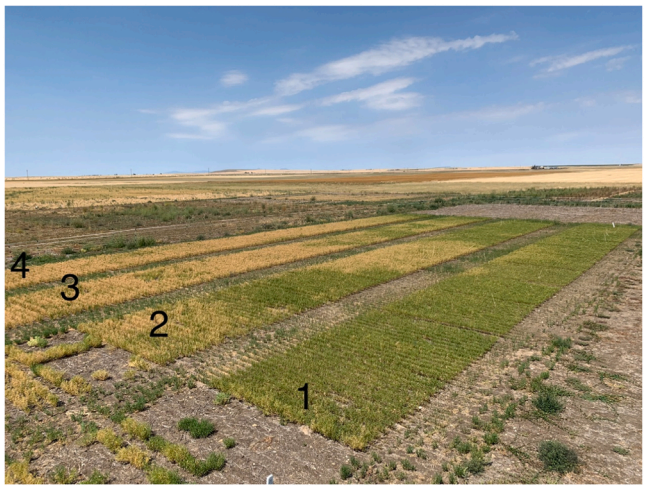


Fig. 1. Spatial heterogeneity in soil depth creates characteristic patterns in crop growth stage of lentil independent of experimental design (randomized complete block). Treatment replicate numbers of the small-plot trial are indicated in black (47°03′20″ N, 109°57′37″ W; 12 August 2021).

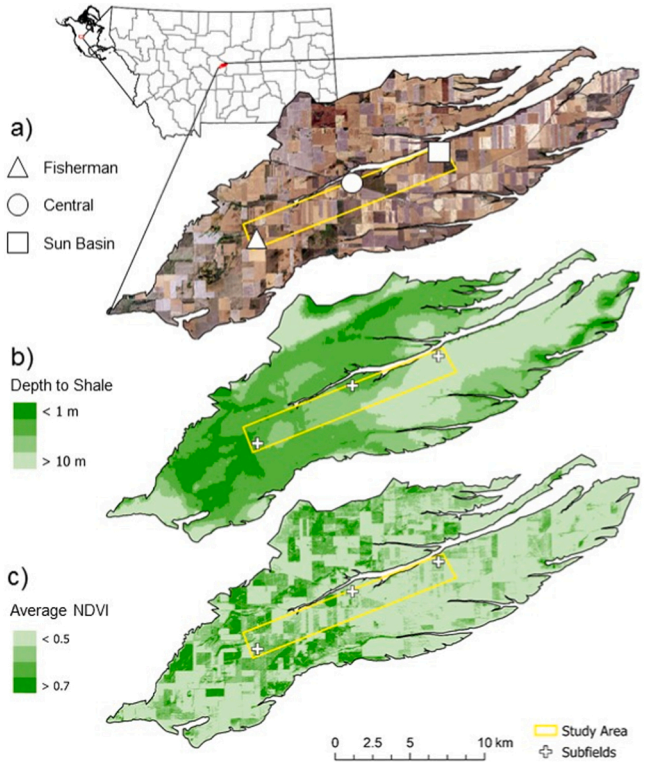


Fig. 2. Moccasin Terrace study area indicating locations of the Fisherman, Central, and Sun Basin study fields within the Judith Basin, Montana. True color aerial image from 2019 (a), shale depth (b); adapted from Sigler et al., 2018), and average NDVI from Sentinel-2 for the 2019 growing season (c).

(Bestwick, 2016; Sigler et al., 2020). Site-specific management such as variable rate seeding and N fertilizer application has strong potential to improve fine-scale N-use efficiency in these environments (Sigler et al., 2020; Hegedus et al., 2022). The use of spatial z_f information in variable rate fertilizer programs can improve net returns and mitigate environmental impacts of high application rates (Bullock and Bullock, 2000; Basso et al., 2011; Vazquez-Amabile et al., 2013). Spatial z_f information can also assist in the interpretation of crop growth, yield, and other data generated by small-plot agronomy experiments initiated without regard for fine-scale heterogeneity in z_f (e.g., Fig. 1).

Government soil databases like those of the United States Natural Resources Conservation Service (USDA, 2020a) contain spatial information regarding the overall distribution of soil texture and horizon development within the soil profile, including depth to subsurface horizons that are dominated by coarse textures and, thus, have limited water storage capacity. However, these resources use discrete polygons and associations of multiple soil series to delineate soil units that do not capture detailed spatial variability of z_f (Tesfa et al., 2009; Sigler et al., 2020). These resources lack intermediate- to fine-scale characterization of z_f and other soil properties (e.g., texture, organic matter content, hydraulic conductivity) that impact crop production (Santanello et al., 2007; Brom et al., 2021), creating the need to characterize spatial heterogeneity by alternative means.

Soil depth can be determined by excavating soil pits, but the methodology is destructive, costly, and labor intensive (Dietrich et al., 1995; Tesfa et al., 2009). Bestwick (2016) used maximum depth of a soil coring tube and block kriging to estimate mean z_f for inclusion as a covariate in assessing crop sequence effects on wheat grain yield and protein concentration in central Montana. Crop yield increased 410 kg ha^{-1} and protein concentration decreased 16 g kg^{-1} for each 10-cm increase in soil depth in the drier of two years in that study. A wheat yield response of this magnitude (41 kg ha^{-1} cm^{-1}) was within the range (10–56 kg ha^{-1} cm^{-1}) reported by Krüger et al. (2020), who detected correlations between yield and z_f ($p < 0.05$) in six study years with below average growing season precipitation, but no correlation ($p = 0.17$) in an additional year when growing season precipitation was above-normal. Importantly, the methods used to characterize soil depth in these studies were time- and resource-intensive and the resulting maps were limited in both extent and resolution.

Sigler et al. (2020) mapped z_f by regressing depth to cobble observations from soil pits with normalized difference vegetation index (NDVI) of barley (*Hordeum vulgare* L.) derived from National Agricultural Imagery Program (NAIP) data acquired during crop senescence (ZGS 90–93; Zadoks et al., 1974). Their results supported the hypothesis that variation in crop greenness is related to water availability as a function of z_f in regions where water limitation affects senescence timing. Images acquired during crop senescence are useful for z_f mapping because thicker soils store more water in the root zone, thus delaying senescence where soils are deeper (Sigler et al., 2020). While NDVI images acquired near mid-senescence (Zadoks 91–92) are likely to correlate with z_f, 'optimal' timing has not been determined.

In fertility assessments, soil depth is typically estimated as depth to refusal of a soil coring tube (e.g., Jones et al., 2011). Soil probing has been used to estimate z_f directly (Williams et al., 2009; Jones et al., 2011; John et al., 2017), in combination with kriging efforts (Bestwick, 2016), and in model validation (Tesfa et al., 2009; Ziadat, 2010; Sigler et al., 2020). Relative to soil pits, the probing method is minimally destructive, providing a simple means of estimating z_f while preserving spectral signatures for future assessments. Cereal crop NDVI of pixel size < 1 m has been shown to correlate positively with z_f within field boundaries (Sigler et al., 2020), but it remains unclear whether correlations between z_f and NDVI exist at larger pixel sizes (1–10 m) and across field boundaries. Our objective was to determine the impact of image spatial resolution on the strength of z_f -NDVI correlations within and across multiple fields in a semi-arid, dryland agroecosystem.

2. Materials and methods

2.1. Study area

The Moccasin Terrace in the Judith River Watershed, MT, United Sates (47°1′52"N, 110°2′14"W - 47°4′58"N, 109°52'39"W; Elevation: 1251–1341 m) is a 260-km^2 gravel-mantled strath terrace with shallow water tables (< 10 m) and shallow soils in terms of surficial fine-textured materials (63 ± 30 cm; μ ± SD; Sigler et al., 2020) overlying

Table 1
Coordinates of soil depth estimates by tile probe with elevation and slope estimates for individual points, grouped by field.

Field	Point	Latitude, Longitude	Tile Probe (cm)	Elevation (m)	Slope (%)
Fisherman	1	47°1′54.6235"N, 110°2′14.2462"W	34	1340	2.7
	2	47°1′54.4822"N, 110°2′14.6485"W	27	1341	2.9
	3	47°1′54.2506"N, 110°2′13.8766"W	36	1339	2.9
	4	47°1′53.8333"N, 110°2′13.9782"W	45	1339	2.9
	5	47°1′53.6374"N, 110°2′15.5339"W	35	1341	1.6
	6	47°1′53.2801"N, 110°2′15.117"W	51	1340	2.7
	7	47°1′53.111"N, 110°2′14.1323"W	40	1339	2.3
	8	47°1′53.0189"N, 110°2′14.6513"W	42	1340	2.7
	9	47°1′52.8333"N, 110°2′13.6198"W	42	1339	2.3
	10	47°1′52.8327"N, 110°2′13.5493"W	39	1339	2.3
	11	47°1′52.7328"N, 110°2′15.0563"W	40	1340	2.7
	12	47°1′52.3277"N, 110°2′14.6873"W	46	1341	0.0
Central	1	47°3′56.6663"N, 109°57'10.7646"W	40	1289	0.0
	2	47°3′56.6132"N, 109°57'11.9894"W	39	1289	0.0
	3	47°3′56.5207"N, 109°57'10.6733"W	89	1289	0.0
	4	47°3′56.1493"N, 109°57'12.2574"W	32	1289	1.4
	5	47°3′56.0262"N, 109°57'11.1066"W	54	1289	1.6
	6	47°3′55.8976"N, 109°57'11.1091"W	70	1289	1.6
	7	47°3′55.6508"N, 109°57'10.5021"W	24	1289	1.6
	8	47°3′55.444"N, 109°57'12.5065"W	34	1289	2.3
	9	47°3′55.2204"N, 109°57'10.9106"W	29	1290	2.3
	10	47°3′54.3955"N, 109°57'12.1505"W	35	1291	2.3
	11	47°3′54.0115"N, 109°57'10.6048"W	34	1291	1.6
	12	47°3′53.8124"N, 109°57'11.6912"W	30	1291	2.3
Sun Basin	1	47°4′58.741"N, 109°52'39.4766"W	34	1251	1.4
	2	47°4′57.9238"N, 109°52'38.1283"W	43	1252	0.0
	3	47°4′57.8421"N, 109°52'39.6367"W	36	1251	1.6
	4	47°4′57.4093"N, 109°52'38.1391"W	35	1251	0.9
	5	47°4′57.0464"N, 109°52'38.853"W	34	1251	0.0
	6	47°4′57.0088"N, 109°52'38.2888"W	42	1251	0.9
	7	47°4′56.7161"N, 109°52'37.9654"W	28	1251	0.9
	8	47°4′56.4443"N, 109°52'39.7602"W	28	1251	0.0
	9	47°4′56.1163"N, 109°52'39.108"W	30	1251	0.0
	10	47°4′56.018"N, 109°52'38.9217"W	26	1251	0.0
	11	47°4′55.9597"N, 109°52'39.535"W	31	1251	0.0
	12	47°4′55.8213"N, 109°52'38.5492"W	25	1251	0.0

calcareous, gravel-dominated subsurface horizons. A 15-km transect of the Moccasin Terrace delineated the study area (Fig. 2). Within this transect, three fields (Fisherman, Central, and Sun Basin) with soils that vary in thickness of fine-textured subsurface horizons (complex of Vertic Argiustolls and Typic Calciustolls, 0–2% slopes; USDA, 2020a) were targeted for z_f characterization. Fields spanned a gradient in depth to an impermeable shale layer which underlies ("perches") the shallow aquifer at depths < 20 m, with implications for field accessibility in spring and precipitation partitioning to overland flow versus deep percolation (Sigler et al., 2018).

Long-term annual precipitation (1909–2020) and air temperature (1911–2020) averaged 390 mm and 6 °C at the Montana State University Central Agricultural Research Center (MSU, 2020) near the center of the study area (Central). Growing season (April – August) precipitation was 75% (2017), 112% (2018), 119% (2019), and 84% (2020) of the long-term average (112-yr; 258 mm) during the study. From 2017–2020, the majority of cropland on the Moccasin Terrace was planted to winter wheat (24–35%), alfalfa (*Medicago sativa* L.; 17–19%), spring wheat (10–22%), barley (6–14%), or left fallow (9–18%; USDA, 2020b).

2.2. Field selection

Multiple ancillary and reference datasets were used in field selection, including CropScape (USDA, 2020b), NAIP (USDA, 2020c), multispectral imagery from an unmanned aerial vehicle (UAV; senseFly eBee, Cheseaux-Lausanne, Switzerland), and a shale surface digital elevation model of the Moccasin Terrace derived from well logs (Sigler et al., 2018). Fields were selected that spanned a known gradient in depth to shale to capture the range of spectrally relevant geologic boundary conditions for water table elevation within the study area. Visual comparisons of the shale surface digital elevation model and growing season average NDVI derived from image reduction of Sentinel-2 (Copernicus, 2020) multispectral data corroborated depth to shale estimates (Fig. 2). Fields were selected where spring wheat was planted in 2020 and winter or spring wheat was planted in 2018 and 2019, no-tillage practices were used, and there was high spectral variability within cereal crop stands observed in 2017 and/or 2015 NAIP imagery. This methodology ensured that at least one viable NDVI map (i.e., a map with application for z_f estimation) could be generated. Moreover, selection occurred so distance among fields was minimized ($\leq$15 km) while the range of predicted shale depth values was maximized (1–17 m). Any potential orographic contributions to weather differences among fields were minimal in the study area given the relative absence of complex topography. Nonetheless, study fields were selected along a transect oriented in the direction of the prevailing wind (i.e., southwest) to minimize potential effects of weather differences among fields.

2.3. Subfield selection and soil depth measurements

A 0.4-ha subfield within each 8–80 ha field was selected for data collection by visually identifying areas of maximum spectral variability that were representative of larger fields, as determined by single-band histogram comparisons. Soil depth within each subfield was randomly sampled at 12 points at Fisherman and Sun Basin, and 42 points at Central. Initial approximations of z_f were made based on the maximum depth penetrated by a tile probe (diam. = 19 mm; Fabian Machine & Welding Inc., Lewistown, MT, United States) at a downward pressure of $>$ 100 MPa. Downward pressure on the probe was estimated by retrofitting a Giddings hydraulic soil probe (15- TS / Model GSRTS; Giddings Machine Company, Windsor, CO, United States) with a pressure gauge.

At Central, a subset of 12 points (of the original 42) were selected at random for z_f sampling three additional times by tile probe (Table S1). The 12 resampled points were withheld, and ordinary kriging with leave-one-out cross validation was performed with the remaining 30 tile probe sampling points (Fig. S1a). To determine if edaphic conditions (e.

Table 2
Parameter estimates for non-linear regression models (Eq. 1; n = 12; df = 9) where average tile probing depth was designated as a predictor and NDVI as a response variable, grouped by sensor, year, acquisition day of year (DOY), and field. Regressions were performed using NDVI of fields planted to spring wheat in central Montana. Highlighted in grey are regressions performed with single-event tile probing depths versus average probing depths.

[a]Sensor	Resolution (m)	Year	DOY	Field	[b]*a*		*b*		*c*	
Sentinel-2	10	2019	193	Central	0.63	* **	0.17	NS	2	NS
			203		0.70	* **	0.15	NS	8	NS
			206		0.69	* **	0.12	.	11	.
			216		0.55	* **	0.09	NS	14	NS
			218		0.46	* **	0.09	NS	13	NS
			233		0.26	* **	0.08	NS	12	NS
		2020	213		0.53	* **	0.09	NS	5	NS
			218		0.48	* **	0.08	NS	7	NS
			228		0.31	* **	0.09	NS	6	NS
UAV	0.1	2017	222		0.31	* **	0.08	.	24	* **
NAIP	0.6		228		0.09	* **	0.11	NS	27	* *
Sentinel-2	10	2020	218	Central	0.47	* **	0.25	NS	19	* **
				Fisherman	0.58	* **	0.11	NS	18	.
				Sun Basin	0.48	* **	0.12	NS	18	* *

Significance Codes: ' * ** ' = $p < 0.001$; ' * * ' = $p < 0.01$; ' * ' = $p < 0.05$; '. ' = $p < 0.1$; 'NS' = $p \geq 0.1$
[a] UAV = unmanned aerial vehicle; NAIP = National Agricultural Imagery Program; Sentinel-2 = Sentinel-2 Satellite System
[b] Parameters a, b, and c represent the upper asymptote, maximum slope, and value of z_f at the maximum slope, respectively

g., soil moisture) at the time of sampling (26 September 2018, 16 October 2018, 30 March 2020, 18 May 2020) biased z_f estimates, analysis of variance was performed with tile probing depth designated as the response variable and sampling event crossed with kriged z_f as a fixed effect. Volumetric water content at 20 cm (VWC_{20}) measured in a cultivated field (continuous spring wheat) within 1 km of the Central subfield at the time of tile probing ranged from 0.27 to 0.43 m^3 m^{-3} across all sampling events (GS3, METER Group, Inc., Pullman, WA, United States; Montana Climate Office, 2022). Ultimately, tile probing depth averages were used as z_f estimates to determine the effects of sensor resolution, acquisition day, and acquisition year on z_f with NDVI relationships at Central. Only one sampling event occurred at Fisherman and Sun Basin, so tile probing depths from this and the corresponding event at Central were used to approximate z_f in multilocation models (Table 1).

2.4. Multispectral sensors and images

Three multispectral sensors (UAV, NAIP, Sentinel-2) of varied spatial resolution (0.1, 0.6, 10 m, respectively) captured images of the subfield at Central during crop senescence in 2017 (UAV, NAIP) and 2019 (Sentinel-2; Copernicus, 2020). The subfield was planted to, and preceded by, hard red spring wheat in both years. The UAV, NAIP, and Sentinel-2 multispectral images were acquired on 10 August 2017, 16 August 2017, and 4 August 2019, respectively. The UAV was an eBee drone (senseFly Ltd., Cheseaux-Lausanne, Switzerland) mounted with a senseFly MultiSpec 4 C sensor (Airinov, Paris, France) with green (550 nm), red (660 nm), red edge (735 nm), and NIR (790 nm) bands. The sensor has an upward-facing irradiance sensor which enables calibration to surface reflectance. SenseFly eMotion 2 software was used for flight planning and execution, and Pix4Dmapper Pro (Pix4D SA, Cheseaux-Lausanne, Switzerland) was used for post-processing derivation of multispectral orthomosaics. Images used to assess impacts of acquisition timing were acquired by Sentinel-2 on 17 July (pre-senescence; Zadoks growth stage $\leq$ 89), 22 July (early senescence; Zadoks = 90), 25 July (early senescence), 6 August (mid-senescence; Zadoks = 91–92), and 21 August 2019 (late senescence; Zadoks = 93). The image used to assess impacts of field was acquired by Sentinel-2 on 5 August 2020 at mid-senescence. An additional image of the Central subfield acquired by Sentinel-2 on 6 August 2019 was compared against that acquired on 5 August 2020 to assess effects of acquisition year.

2.5. Statistical analysis

Average tile probing depth values of the 12 withheld probing points at Central were regressed against kriged estimates of z_f, as well as NDVI from UAV (10 August 2017), NAIP (16 August 2017), and Sentinel-2 (4 August 2019) sensors. Moran's I detected no autocorrelation in any model, and assumptions of normality and homoscedasticity were satisfied by Jarque-Bera and Breusch Pagan tests, respectively. Still, visual inspections of fitted values versus residuals raised concerns of heteroscedasticity, so nonlinear regression was performed using the nls function of the package nlme (Pinheiro et al., 2021) in R statistical software (R Core Team, 2021). Soil depth was designated as the predictor variable and NDVI as the response variable in a logistic model:

$$NDVI = \frac{a}{1 + e^{-b(z_f - c)}} \quad (1)$$

where parameters a, b, and c represent the upper asymptote, maximum slope, and value of z_f at the maximum slope, respectively. Parameter estimates and significance for select converged models are indicated in Table 2. Causality and precedence (Sigler et al., 2020) were considered when designating z_f as the predictor variable and NDVI as the response, although it is recognized that these variables will be reversed in practical applications.

Linear regression was used to determine whether z_f predicted NDVI independent of acquisition day, sensor, field, and acquisition year. However, assumptions of normality and homoscedasticity were often violated when using linear regression, and transformations showed inconsistencies among image groups in resolving these violations. Thus, when evaluating interactions, observations > 50 cm were discarded to capture only the linear portion of the curve, thereby eliminating severe violations of normality and homoscedasticity ($p < 0.01$) in linear models. Precedence (Sigler et al., 2020) and apparent bimodality in average tile probe depths with the approximate minimum frequency dividing the modes at 50 cm (Fig. S2) gave justification for this approach. For these assessments, analysis of variance was performed in base R (R Core Team, 2021) with z_f designated as a fixed effect crossed with either date, sensor, field, or year in the interaction term.

2.6. Imagery based soil depth models

Tile probing depth from a single sampling event was designated as the response variable in multilocation models assessing the ability of Sentinel-2 data to predict soil depth across management boundaries. Two separate modelling approaches relying on image reduction

Table 3
Summary statistics of linear models assessing effects of average (Central; n = 6; df = 4) or single event (All; n = 32; df = 30) tile probing depths (< 50 cm) on NDVI from multiple sensors acquired in different years, on different days of the year (DOY), and at different growth stages (Zadoks et al., 1974).

Sensor	Field(s)	Year	DOY	Zadoks	*p*-value	RMSE	R^2
Sentinel-2	Central	2019	206	90	0.018	0.03	0.74
NAIP	Central	2017	228	93	0.134	0.02	0.34
UAV	Central	2017	222	92	0.061	0.03	0.53
Sentinel-2	All	2020	218	91–92	< 0.001	0.05	0.54

techniques were used to predict tile probing depth from categorical (class) and continuous ($\hat{z}_f$) estimates of z_f. In the categorical approach, supervised classification (Classification and Regression Trees classifier; CART; Breiman et al., 1984) of 4-band (R, G, B, NIR) Sentinel-2 imagery (10 m resolution) was carried out in Google Earth Engine (Goerlick et al., 2017). Average seasonal (16 June through 4 September 2020) reflectance values of each band were used as training and validation data in the CART classification. NDVI from NAIP (0.6 m resolution) was used as ancillary data in the delineation of training and validation polygons, within which every Sentinel-2 pixel was treated as a training or validation point. Training polygons included a total of 124 and 126 pixels for 'deep' and 'shallow' classes, respectively, satisfying the 100-pixel guideline suggested by Campbell & Wynne (2011).

In the continuous approach, mean ($NDVI_{MEAN}$) and max NDVI ($NDVI_{MAX}$) for the period 16 June to 04 September 2020 were included

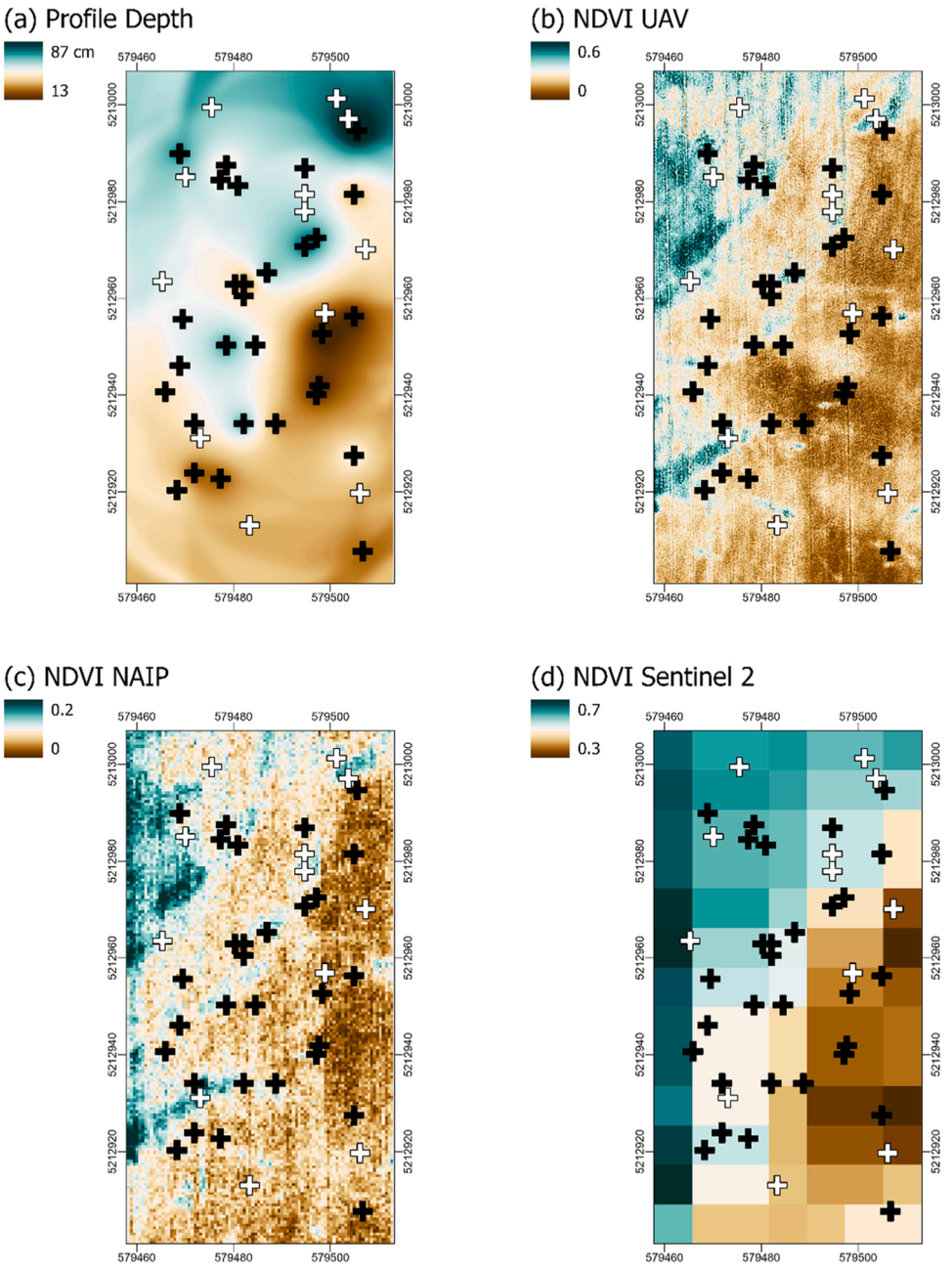


Fig. 3. Kriged estimates of soil depth based on probing (a) and NDVI from different imagery sources (b-d). NDVI was acquired by an unmanned aerial vehicle (UAV; b), the National Agricultural Imagery Program (NAIP; c), and Sentinel-2 satellite systems (Sentinel 2; d) on 10 August 2017, 16 August 2017, and 4 August 2019, respectively. Black symbols are the 30 initial tile probing points and white symbols are the 12 points withheld for cross validation and NDVI regressions.

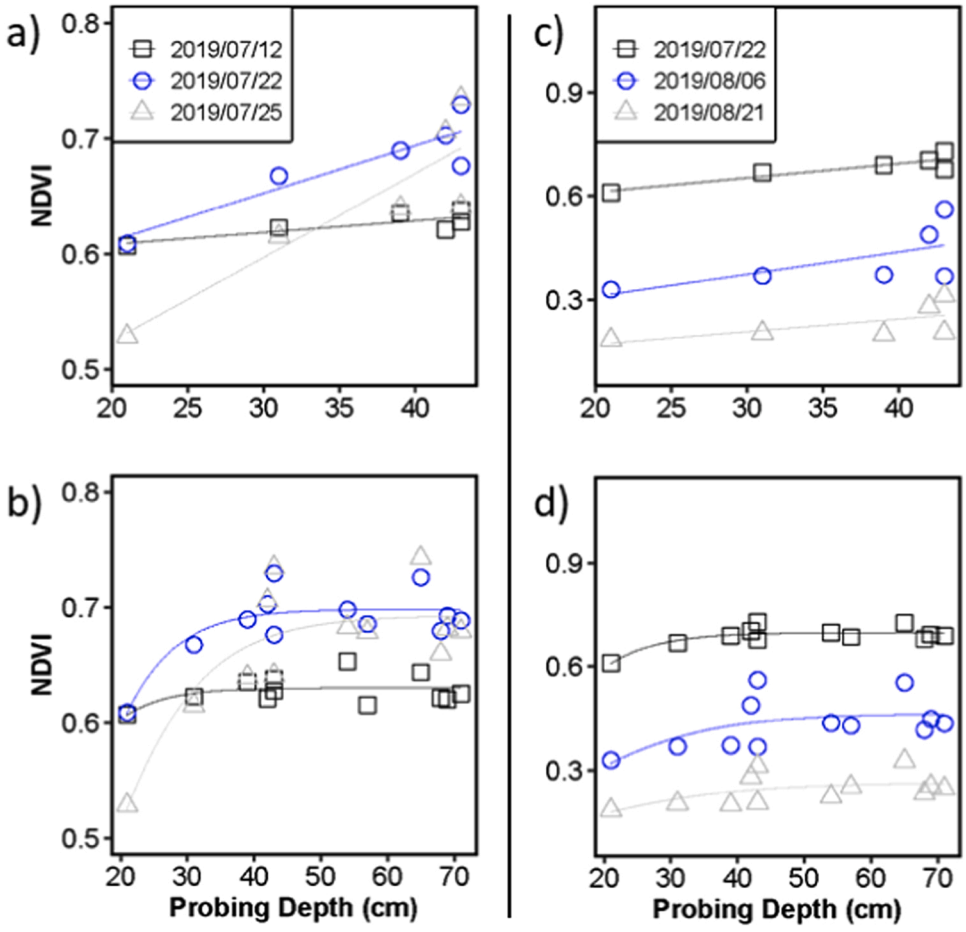


Fig. 4. Impacts of image acquisition timing on linear (a, c) and non-linear (b,d) regression models for probing depth and NDVI from Sentinel-2. Probing depth observations > 50 cm were excluded in linear regressions (a, c). Among pre- (12 July 2019) and early- (22 July, 25 July 2019) senescence images (a, b), image date was important for determining the slope (linear) and shape (non-linear) of probing depth-NDVI regression lines. Impacts of acquisition date on probing depth-NDVI relationships were less pronounced among early- (22 July 2019), mid- (6 August 2019), and late- (21 August 2019) senescence images (c,d).

as independent variables in a logistic model:

$$\hat{z}_f = \frac{-\log\left(\frac{NDVI_{MAX}}{NDVI_{MEAN}} - 1\right) - ab}{a} \quad (2)$$

where parameters a and b are empirically derived constants set to 0.013 and − 22.3, respectively, based on an iterative parameterization that optimized the regression statistics of predicted $\hat{z}_f$ and observed z_f (tile probing depth).

Regression statistics of linear models designating categorical (class) or continuous ($\hat{z}_f$) estimates of z_f as fixed effects were compared to those of the model designating the 'best' Sentinel-2 NDVI image (i.e., the image explaining the most variability in tile probing depth across fields) as a fixed effect. Initially, tile probing depth was designated as the response variable with class, $\hat{z}_f$, or NDVI designated as fixed effects crossed with field (class × field, $\hat{z}_f$ × field, or NDVI × field). Because interactions (at $p < 0.05$) were not observed in any model, both field and the interaction term were dropped as fixed effects and the ability of z_f proxies (class, $\hat{z}_f$, NDVI) to explain variance in z_f observations (tile probing depth) across fields was assessed with linear modelling in base R. To assess year-dependencies of z_f proxies, mixed-effects modelling was performed where NDVI was treated as the response variable and acquisition year crossed with tile probing depth, $\hat{z}_f$, or class was designated as a fixed effect. Field was designated as a random effect in these models, allowing comparisons of conditional coefficients of determination (conditional R^2), representing variability explained by the entire model, with marginal coefficients of determination (marginal R^2), representing variability explained by fixed effects only.

3. Results and discussion

3.1. Kriged estimates of soil depth

Ordinary kriging based on a single tile probe sampling event (n = 30) with optimized variogram statistics (lag size; = 4.6; nugget = 0; range = 37; sill = 332; Fig. S1a) performed modestly well based on leave-one-out cross validation (Fig. S1b) and regression of predicted z_f against average tile probing depths at the 12 withheld points (y = 0.82x + 9; RMSE = 12 cm; $R^2 = 0.50$; $p = 0.006$). The fitted spherical variogram in Bestwick (2016) did not describe the empirical variogram in the current study, but grid sampling of z_f in the former study versus fully randomized sampling in the latter may help explain this difference, given that distances between sampling points were not consistent. Furthermore, parameters used to describe the variogram fit in the current study (nugget = 0; range = 37; sill = 332) were outside the range of parameters (nugget = 149–266; range = 31–34; sill = 111–187) reported by Bestwick (2016), although this is not surprising given that different variograms were selected (stable versus spherical).

3.2. Probing depth and NDVI

Average tile probing depth at Central explained 53%, 34%, and 74% of the variability in NDVI from UAV, NAIP, and Sentinel-2 sensors, respectively (Table 3), compared to 50% of the variability in kriged estimates of z_f (see 'Kriged Estimates of Soil Depth'). Coefficients of determination from these linear regressions (Table 3; R^2 = 0.34–0.74) were within the range (R^2 = 0.32–0.89) of those reported in previous studies investigating relationships between soil depth and spectral indices (R^2 = 0.41, McKenzie and Ryan, 1999; R^2 = 0.32–0.89, Taylor

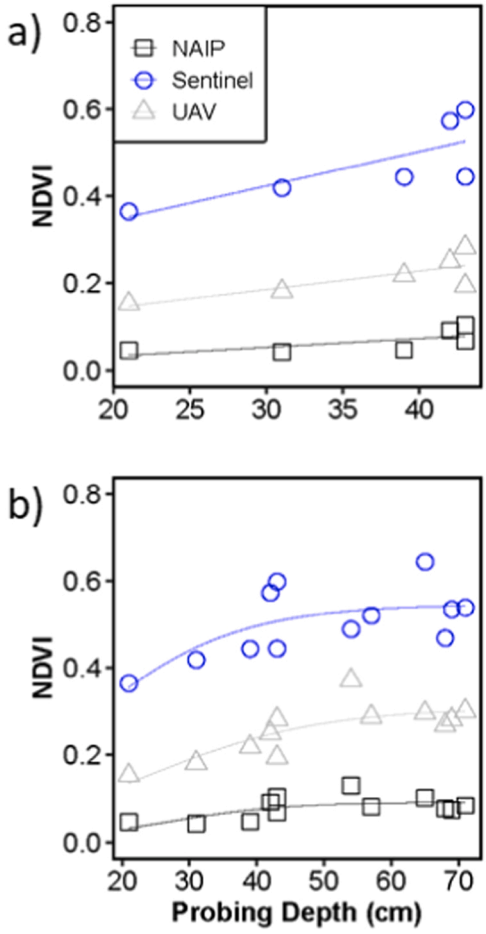


Fig. 5. Impacts of sensor on linear (a) and non-linear (b) regression models for probing depth versus NDVI acquired by an unmanned aerial vehicle (UAV), the National Agricultural Imagery Program (NAIP), and Sentinel-2 satellite systems (Sentinel) on 10 August 2017, 16 August 2017, and 4 August 2019, respectively. Probing depth observations > 50 cm were excluded in linear regressions. Impacts of sensor on the slope (linear) and shape (non-linear) of probing depth-NDVI regression lines were minimal, while impacts to the magnitude of NDVI were likely associated with differences in acquisition date.

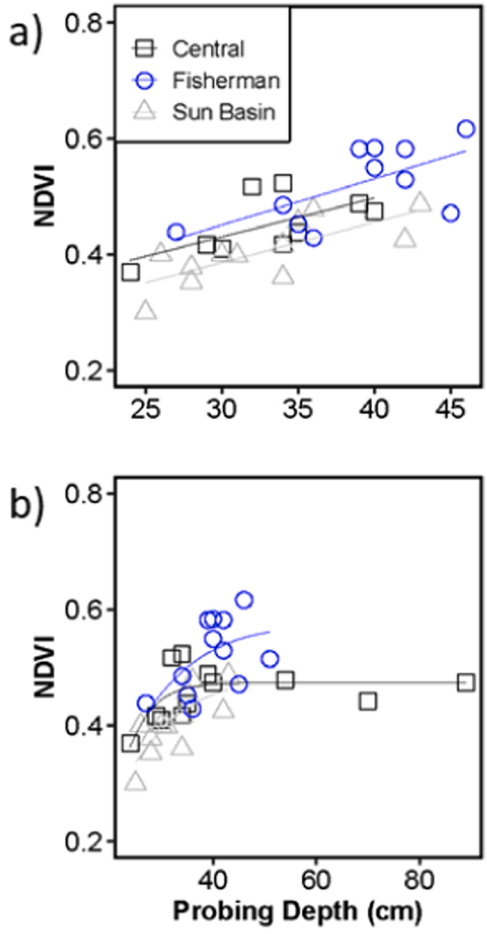


Fig. 6. Impacts of field on linear (a) and non-linear (b) regression models for probing depth versus NDVI acquired by Sentinel-2 satellite systems on 5 August 2020. Probing depth observations > 50 cm were excluded in linear regressions. Impacts of field on the slope and y-intercept (linear) and the shape (non-linear) of probing depth-NDVI regression lines were minimal.

et al., 2013; $R^2 = 0.46$, Sigler et al., 2020). When NDVI from the NAIP sensor was analyzed separately, the relationship was not statistically significant ($p = 0.134$). Otherwise, regression statistics indicated imagery of fine (0.1 m; UAV; $p = 0.061$) and intermediate (10 m; Sentinel-2; $p = 0.018$) resolution adequately characterized heterogeneity in z_f, supported by visual comparisons of kriged z_f and multi-sensor NDVI (Fig. 3). The probing depth with NDVI relationship was highly dependent on acquisition day ($p = 0.01$; Fig. 4a) when comparing images acquired pre- and early-senescence (ZGS 89–91). However, probing depth predicted NDVI independent of acquisition day ($p = 0.74$; Fig. 4c), sensor ($p = 0.22$; Fig. 5a), field ($p = 0.94$; Fig. 6a), and year ($p = 0.68$; Fig. 7a) when comparing images acquired during crop senescence (ZGS 90–93).

Sensor resolution could not be assessed independently of acquisition day, but all sensors acquired images during crop senescence when acquisition day is least likely to affect the slope of z_f -NDVI regression lines and most likely to affect the y-intercept (Fig. 4c). Therefore, differences in the slopes of z_f -NDVI regression lines among sensors (Fig. 5a) more likely reflected differences in spatial resolution than differences in y-intercept. Since no statistical $z_f \times$ sensor interaction was observed ($p = 0.22$), and since the weakest z_f-NDVI correlation (NAIP; $R^2 = 0.34$; 0.6 m) did not correspond with the coarsest (Sentinel-2; 10 m) or finest (UAV; 0.1 m) resolution sensor, it was concluded that correlations between z_f and NDVI do not depend on sensor, as long as the image resolution is $\leq$ 10 m and images are acquired during crop senescence. Similarly, because acquisition day was held constant in assessments of the impacts of field and year on z_f -NDVI relationships, yet no z_f interactions with field ($p = 0.94$; Fig. 6) or year ($p = 0.68$; Fig. 7) were detected, it was concluded that correlations between z_f and NDVI depend less on field and acquisition year than on acquisition day.

Soil depth-NDVI relationships seem to depend more on seasonal (i.e., intra-annual) variation in remote sensing spectra than on inter-annual variability, field, and spatial resolutions $\leq$ 10 m. These results add to a growing body of evidence highlighting the value of multi-temporal remote sensing datasets in characterizing spatially variable soil properties (e.g., Maynard and Levi, 2017). Images acquired at different crop stages may be used to spatially characterize different soil properties. For example, Webb et al. (2021) found that early-season NDVI explained more variability in surface pH than late season NDVI in a field in southwest Montana, while the current study demonstrates that late-season NDVI explains more variability in z_f than early season NDVI within fields. Future research should investigate temporal sensitivity of NDVI-z_f relationships in different geographic regions and climates.

3.3. Multilocation soil depth estimates

Soil depth modelling across fields showed that the Sentinel-2 sensor's 'best' multilocation z_f-NDVI relationship (5 August 2020; Table 4) outperformed z_f proxies relying on image reduction techniques.

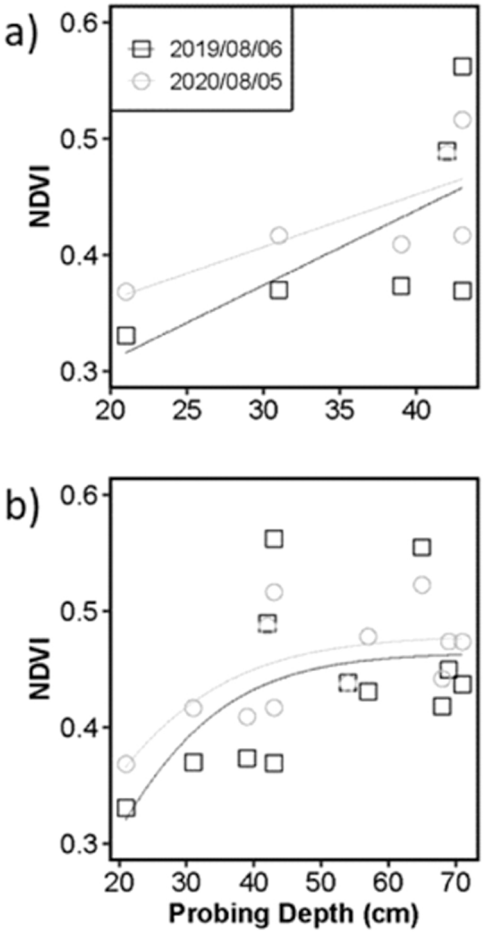


Fig. 7. Impacts of acquisition year on linear (a) and non-linear (b) regression models for probing depth versus NDVI acquired by Sentinel-2 satellite systems on 6 August 2019 and 5 August 2020. Probing depth observations > 50 cm were excluded in linear regressions. Impacts of year on the slope (linear) and shape (non-linear) of probing depth-NDVI regression lines were minimal.

Table 4
Multilocation linear models (n = 32; df = 30) assessing effects of single event probing depth (<50 cm) on NDVI (Sentinel-2; 5 August 2020), categorical soil depth estimates (class), and continuous soil depth estimates ($\hat{z}_f$).

Fixed Effect	*p*-value	RMSE (cm)	R^2
NDVI	< 0.001	4	0.54
class	0.003	5	0.24
$\hat{z}_f$	0.001	5	0.30

Specifically, the empirical spectral index ($\hat{z}_f$; Eq. 2) and the CART classifier (class) explained 30% and 24% of the variability in tile probing depth, compared to 54% of the variability explained by well-timed Sentinel-2 NDVI. The continuous model outperformed the categorical model based on coefficients of determination ($R^2 = 0.30$ vs 0.24) and *p*-values ($p = 0.001$ vs 0.003), but root mean square errors were equivalent (5 cm). Despite coefficients of determination for class and $\hat{z}_f$ predictions falling outside the range of previously published z_f predictions ($R^2 = 0.32$–0.89), *p*-values and root mean square errors were quite small for both models. Furthermore, the ability of class to predict NDVI did not depend on acquisition year ($p = 0.12$), but the acquisition year interaction was stronger ($p = 0.053$–0.062) for both $\hat{z}_f$ and tile probing depth models (Table S3), suggesting predictions of the CART classifier (class) may be relatively tolerant to inter-annual variability in remote sensing spectra.

While the error matrix and accuracy statistics of the CART classification (Table S4) were not directly applicable to z_f characterization (i.e., z_f was not measured in all training/test areas) results indicate that spatial patterns of NDVI in well-timed NAIP images are reproducible with Sentinel-2 imagery on a landform where such patterns are known to correlate with z_f. Visual agreement between classified Sentinel-2 data and NDVI from well-timed UAV imagery at Central suggests the categorical approach did a reasonable job of delineating z_f features (Fig. S3). While previous research characterized soil depth at high spatial resolutions (< 10 m; Bestwick, 2016; Sigler et al., 2020), the current study improved the extent of z_f prediction three-fold (145 vs 44 ha) by including multiple fields. Results support the use of multitemporal imagery of intermediate resolution for characterizing z_f across fields.

A field effect on NDVI was detected in a well-timed Sentinel-2 NDVI image ($p < 0.001$; 5 August 2020; Table 4) in the current study, but a $z_f \times$ field interaction was not ($p = 0.94$). Since fields were selected to capture a known gradient in depth to a subsurface shale layer (Fig. 2), and since the highest NDVI values corresponded with the shallowest shale contacts (Fig. 2; Fig. 6), it is suggested that fine-scale spatial patterns in NDVI were largely driven by z_f but that geologic differences exerted a regional influence on these patterns. The similarity in slopes of z_f-NDVI relationships with field suggests that edaphic (z_f) and geologic (shale depth) controls on water availability did not interact but had additive effects on NDVI of wheat. Thus, well-timed NDVI (Zadoks 90–93) derived for multiple fields in the study area might be best conceptualized as an *effective* z_f index, reflecting differences in water supply related to variability in z_f but also to other factors. It is theorized that crop senescence occurred earlier where shale contacts were deeper because plant roots were unable to access shallow groundwater in these locations, eliminating the 'sub-irrigation' effect observed at shallower shale contacts. This is consistent with previous research (Sigler et al., 2018) which demonstrated that depth to shale altered hydrologic partitioning to deep percolation and, by extension, plant uptake.

Results of the present study stress the importance of image timing for NDVI-based predictions of z_f across fields, as a single well-timed NDVI image explained more variability in tile probing depth than z_f predictions derived from image reduction and machine learning techniques. Results also highlight the inherent value of spaceborne sensors given their large geographical extent and frequent revisit times (Gholizadeh et al., 2018). Where satellite data are not available, well-timed UAV-based imaging is an alternative.

Remote observation of z_f offers a promising means of assessing vulnerability to nutrient loss in agricultural systems, given that this soil property can impact N leaching and denitrification loss (John et al., 2017; Sigler et al., 2018, 2020, 2022), as well as mineralization of soil organic N during fallow (John et al., 2017; Sigler et al., 2022) and in the offseason (Fordyce, 2022). These relationships have relevance for site-specific management strategies that could use spectral indices to optimize N applications and balance potential revenue to possible loss of inputs (Preza-Fontes et al., 2019; Sigler et al., 2020; Zhao et al., 2020; Argento et al., 2021; Hegedus et al., 2023). The analysis allows producers to direct N fertilizer to areas where it will have the greatest impact on crops and least impact on the environment, and together with site-specific assessment of fertilizer response, can inform more cost-effective management strategies.

4. Conclusion

Linear and non-linear regression of probing depth and NDVI from three sources indicated z_f -NDVI relationships were not dependent on sensor resolution or field boundary when images were acquired during senescence. Fine (<1 m) and intermediate (≤ 10 m) resolution NDVI acquired during crop senescence predicted z_f with equivalent or greater accuracy than the time- and resource-intensive sampling efforts involved in kriging. The research highlights the potential for satellite- and secondarily UAV-based imagery to characterize spatial

heterogeneity of z_f—and more generally, soil water supply—across field boundaries and improves our understanding of the practical constraints on spectral and other z_f proxies. There is strong potential for imagery-based z_f indicators to provide site-specific decision support in N fertility management. Future work should use z_f indices to develop specific practices which balance potential revenues and environmental loss.

Funding

This work was supported by the USDA WSARE program grant no. GW18–179 and Hatch grant no. MONB00499/ project accession no. 1022481 from the USDA National Institute of Food and Agriculture.

CRediT authorship contribution statement

SIF, PMC, CJ, JOE, WAS, SE, and SLP contributed to data interpretation and revision and editing of the article. SIF was responsible for data collection, analysis, and interpretation, and initially drafted the article.

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.

Data Availability

Data will be made available on request.

Acknowledgments

The authors would like to thank the cooperating farmers and technicians who made this work possible. Special thanks to Sally Dahlhausen, who assisted with field data collection and served as the lead lab technician for the project.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.agwat.2023.108524.

References

Argento, F., Anken, T., Abt, F., Vogelsanger, E., Walter, A., Liebisch, F., 2021. Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. Precis. Agric. 22 (2), 364–386. https://doi.org/10.1007/s11119-020-09733-3.

Basso, B., Ritchie, J.T., Cammarano, D., Sartori, L., 2011. A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. Eur. J. Agron. 35 (4), 215–222. https://doi.org/10.1016/j.eja.2011.06.004.

Bestwick, M.L., 2016. Water stress in Montana cropping systems: effects of cultivar, management, and environment on crop production in dryland systems (Master's thesis, Montana State University-Bozeman, College of Agriculture).

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees. Wadsworth Int. Group 37 (15), 237–251. https://doi.org/10.1201/9781315139470.

Brom, J., Duffková, R., Haberle, J., Zajíček, A., Nedbal, V., Bernasová, T., Křováková, K., 2021. Identification of infiltration features and hydraulic properties of soils based on crop water stress derived from remotely sensed data. Remote Sens 13 (20), 4127. https://doi.org/10.3390/rs13204127.

Bullock, D.S., Bullock, D.G., 2000. From agronomic research to farm management guidelines: a primer on the economics of information and precision technology. Precis. Agric. 2 (1), 71–101. https://doi.org/10.1023/A:1009988617622.

Campbell, J.B., Wynne, R.H., 2011. Introduction to remote sensing. Guilford Press. https://doi.org/10.3390/rs5010282.

Carr, P., Fordyce, S.I., Lachowiec, J., Bishop, S., Fryer, H.K., Dahlhausen, S., 2021. Replacing fallow with warm-season annual crops in dryland wheat systems. Agron. J. 113 (5), 4142–4160. https://doi.org/10.1002/agj2.20809.

Copernicus Sentinel-2 data, 2020. Processed by European Space Agency.

Dietrich, W.E., Reiss, R., Hsu, M.L., Montgomery, D.R., 1995. A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. Hydrol. Process. 9 (3–4), 383–400. https://doi.org/10.1002/hyp.3360090311.

Fordyce, S.I., 2022. Spatial patterns in soil depth and implications for offseason nitrogen dynamics in dryland wheat systems of central Montana. (Master's thesis, Montana State University-Bozeman, College of Agriculture).

Gholizadeh, A., Žižala, D., Saberioon, M., Borůvka, L., 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel 2 spectral imaging. Remote Sens. Environ. 218, 89–103. https://doi.org/10.1016/j.rse.2018.09.015.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Grylls, J.P., Webb, J., Dyer, C.J., 1997. Seasonal variation in response of winter cereals to nitrogen fertilizer and apparent recovery of fertilizer nitrogen on chalk soils in southern England. J. Agric. Sci. 128 (3), 251–262. https://doi.org/10.1017/S0021859697004206.

Hegedus, P.B., Ewing, S., Jones, C., & Maxwell, B.D., 2023. Using spatially variable nitrogen application and crop responses to evaluate crop nitrogen use efficiency. Nutr. Cycl. Agroecosyst., https://doi.org/10.21203/rs.3.rs-2126328/v1.

Hegedus, P.B., Maxwell, B.D., Mieno, T., 2022. Assessing performance of empirical models for forecasting crop responses to variable fertilizer rates using on-farm precision experimentation. Precis. Agric. https://doi.org/10.1007/s11119-022-09968-2.

John, A.A., Jones, C.A., Ewing, S.A., Sigler, W.A., Bekkerman, A., Miller, P.R., 2017. Fallow replacement and alternative nitrogen management for reducing nitrate leaching in a semiarid region. Nutr. Cycl. Agroecosyst. 108 (3), 279–296. https://doi.org/10.1007/s10705-017-9855-9.

Jones, C., Chen, C., McVay, K., Stougaard, B., Westcott, M., Eckhoff, J., Greenwood, M., 2011. Measured and predicted temporal changes in soil nitrate-N levels from late summer to early spring in Montana (March). Proc. West. Nutr. Manag. Conf. Vol. 9, 77–82. https://www.westernnutrientmanagement.org/proceedings/?action=years.

Krüger, H., Frolla, F., Zilio, J., 2020. Soil depth, water availability and wheat yield in the austral pampa of Argentina. Cienc. Del. Suelo 38 (2), 274–283. https://www.ojs.suelos.org.ar/index.php/cds/article/view/595.

Maynard, J.J., Levi, M.R., 2017. Hyper-temporal remote sensing for digital soil mapping: characterizing soil-vegetation response to climatic variability. Geoderma 285, 94–109. https://doi.org/10.1016/j.geoderma.2016.09.024.

McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89 (1–2), 67–94. https://doi.org/10.1016/S0016-7061(98)00137-2.

Montana Climate Office, 2022. University of Montana Mesonet Data Downloader. http://shiny.cfc.umt.edu:3838/mesonet-download/ (Accessed March 2022).

MSU, 2020. Central Agricultural Research Center, Moccasin, MT. Weather Data Viewer. https://agresearch.montana.edu/carc/weather/2020/2020%20Crop%20Year%20Thru%20August.pdf (Accessed March 2022).

Pinheiro J., Bates D., DebRoy S., Sarkar D., R Core Team, 2021. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–153, https://CRAN.R-project.org/package=nlme.

Preza-Fontes, G., Bhattarai, R., Christianson, L.E., Pittelkow, C.M., 2019. Combining environmental monitoring and remote sensing technologies to evaluate cropping system nitrogen dynamics at the field-scale. Front. Sustain. Food Syst. 3, 8 https://doi.org/10.3389/fsufs.2019.00008.

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Santanello Jr, J.A., Peters-Lidard, C.D., Garcia, M.E., Mocko, D.M., Tischler, M.A., Moran, M.S., Thoma, D.P., 2007. Using remotely-sensed estimates of soil moisture to infer soil texture and hydraulic properties across a semi-arid watershed. Remote Sens. Environ. 110 (1), 79–97. https://doi.org/10.1016/j.rse.2007.02.007.

Sigler, W.A., Ewing, S.A., Jones, C.A., Payn, R.A., Brookshire, E.J., Klassen, J.K., Weissmann, G.S., 2018. Connections among soil, ground, and surface water chemistries characterize nitrogen loss from an agricultural landscape in the upper Missouri River basin. J. Hydrol. 556, 247–261. https://doi.org/10.1016/j.jhydrol.2017.10.018.

Sigler, W.A., Ewing, S.A., Jones, C.A., Payn, R.A., Miller, P., Maneta, M., 2020. Water and nitrate loss from dryland agricultural soils is controlled by management, soils, and weather. Agric. Ecosyst. Environ. 304, 107158 https://doi.org/10.1016/j.agee.2020.107158.

Sigler, W.A., Ewing, S.A., Wankel, S.D., Jones, C.A., Leuthold, S., Brookshire, E.N.J., Payn, R.A., 2022. Isotopic signals in an agricultural watershed suggest denitrification is locally intensive in riparian areas but extensive in upland soils. Biogeochem. 158 (2), 251–268. https://doi.org/10.1007/s10533-022-00898-9.

Taylor, J.A., Jacob, F., Galleguillos, M., Prevot, L., Guix, N., Lagacherie, P., 2013. The utility of remotely-sensed vegetative and terrain covariates at different spatial resolutions in modelling soil and water table depth (for digital soil mapping). Geoderma 193, 83–93. https://doi.org/10.1016/j.geoderma.2012.09.009.

Tesfa, T.K., Tarboton, D.G., Chandler, D.G., McNamara, J.P., 2009. Modeling soil depth from topographic and land cover attributes. Water Resour. Res. 45 (10) https://doi.org/10.1029/2008WR007474.

USDA, 2020a. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at https://sdmdataaccess.sc.egov.usda.gov. (Accessed October 2020).

USDA, 2020b. National Agricultural Statistics Service Cropland Data Layer (Online) http://nassgeodata.gmu.edu/CropScape/. (Accessed October 2020).

USDA, 2020c. Farm Service Agency NAIP Imagery https://www.fsa.usda.gov/programs-and-services/aerial-photography/imageryprograms/naip-imagery/. (Accessed October 2020).

Vazquez-Amabile, G.G., Gonzalo, M., Pella, M., Galbusera, S., Cueto, G.B., 2013, July. Evaluation of the variable rate fertilization in winter crops for shallow soils using soil depth maps and crop simulation models in southeastern Pampas, Argentina. In American Society of Agricultural and Biological Engineers Annual International Meeting. Kansas City, Missouri.

Vomocil, J.A., Ramig, R.E., 1984. Annual cropping on shallow soils (STEEP Bulletin SR701). Cooperative Extension Service, Oregon State University, Corvallis.

Webb, H., Barnes, N., Powell, S., Jones, C., 2021. Does drone remote sensing accurately estimate soil pH in a spring wheat field in southwest Montana? Precis. Agric. 22 (6), 1803–1815. https://doi.org/10.1007/s11119-021-09812-z.

Williams, C.J., McNamara, J.P., Chandler, D.G., 2009. Controls on the temporal and spatial variability of soil moisture in a mountainous landscape: the signature of snow and complex terrain. Hydrol. Earth Syst. Sci. 13 (7), 1325–1336. https://doi.org/10.5194/hess-13-1325-2009.

Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Weed Res 14 (6), 415–421. https://doi.org/10.1111/j.1365-3180.1974.tb01084.x.

Zhao, J., De Notaris, C., Olesen, J.E., 2020. Autumn-based vegetation indices for estimating nitrate leaching during autumn and winter in arable cropping systems. Agric., Ecosyst. Environ. 290, 106786 https://doi.org/10.1016/j.agee.2019.106786.

Ziadat, F.M., 2010. Prediction of soil depth from digital terrain data by integrating statistical and visual approaches. Pedosphere 20 (3), 361–367. https://doi.org/10.1016/S1002-0160(10)60025-2.