Introduction

Dairy cropping systems and concentrated animal feeding operations (CAFOs) are farm enterprises that convert on-farm grown and purchased feeds into milk and/or meat and manure (Haith & Atkinson, 1977). Manure is used to fertilize crops, and good manure management follows the four R's of fertilizer application (Right rate, source, placement, and timing) to minimize nutrient loss to surface and subsurface runoff (Easton et al., 2017). For example, in cold-climate agroecosystems, winter-time manure application is banned because frozen soils increase nutrient runoff potential. Also, farmers must pay close attention to the weather during the summer as to not apply manure too soon before a major rainfall event. Due to a variety of factors (e.g. inaccurate weather forecasts, limited manure storage) the four R's are not always followed to a tee, thus accurate water quality models are needed to estimate the origin of nutrient pollution for environmental decision making (Fu et al., 2020). Even for regulated CAFOs that require nutrient management plans and are told which fields are acceptable for manure application, farms and state agencies are not required to disclose manure application dates if they were recorded (Shea et al., 2022). Less is known about smaller, unregulated CAFOs. Thus, there is a gap in knowledge surrounding the location and timing of manure application events. Remote sensing can be used to estimate the spatial and temporal nature of manure applications, providing modelers with an important explanatory variable of nutrient loss (Chugg et al., 2022; Dodin et al., 2021; Shea et al., 2022).

So far, only three remote sensing based methods for manure application detection have been published. Radar data was used to quantify changes in soil moisture to detect liquid manure application (Shea et al., 2022). A convolution neural network was trained using labeled imagery to find changes in image time series, which was used to detect manure application events in real time (Chugg et al., 2022). The final method uses multispectral band indices, taking advantage of the unique spectral reflectance of manure (Dodin et al., 2021).

This project uses the multispectral imagery approach because it is the simplest of three methods currently researched for manure application event detection. The research using multi-spectral imagery for detecting manure applications uses spectral band indices originally developed to estimate crop residues and soil organic carbon (SOC) content. The indices for estimating SOC content were used to detect exogenous organic matter (EOM, analogous to manure application) applications due to similarities in the chemical composition of SOC and EOM. The goal of this research is to replicate the methods of (Dodin et al., 2021) at a dairy farm in Vermont.

Methods

Datasets and Data Processing

Ground observations - Predictand

Actual manure application dates for two fields on a farm in west-central Vermont were recorded: manure was applied on the first field, AHS, on 10/6 of 2018, 2019, and 2020, and at the second field, DC, manure was applied on 10/12, 9/28, and 10/1 of 2019, 2020, and 2021. The dates of the crop harvest (corn silage, all years) for each year were also recorded. These dates were recorded from on-field visits and personal communication with the farmer.

Remote Sensing Data - Predictor

Sentinel 2 multispectral image collections (S2) were obtained over the fields for the known manure application dates using google earth engine (GEE) python api. Image collections were obtained for each year, each site, and each period (pre and post application). The GEE developer workflow for the s2cloudless mask was used to remove cloud and shadow pixels (jdbcode, 2022). After removing these pixels, the median of the collections were used to represent the fields' pixels for the period. For pre application, the date range from crop harvest to the day before application was used, and for post application, the date range from the day after application to an arbitrary but reasonable (no more than 15) number of days after application was used (Table 1). Some field + year combinations were not able to be used in the analysis due to too many cloud and shadow pixels. Usability of each field + year combination was examined by displaying the pre and post period median images of the image collections.

Table 1 - Date ranges for per and post manure application periods. Note there were no S2 images available in 2018.

Site	Year	Pre date range	Post date range	Used? If not because of (pre, post, or both)
AHS	2018	NA	NA	No (both)
AHS	2019	09/26 - 10/05	10/07 - 10/22	No (pre)
DC	2019	09/29 - 10/11	10/13 - 10/28	Yes
AHS	2020	09/19 - 10/05	10/07 - 10/22	Yes
DC	2020	09/13 - 09/27	09/29 - 10/14	Yes
DC	2021	09/24 - 09/30	10/03 - 10/17	No (both)

Pre-Processing and Processing Workflow

This section serves as a pseudocode for the entire pre-processing and processing workflow:

- Identify the area of interests (aoi) (fields) as geojson code using using website that generates the code - https://geojson.io/#map=2/0/20
- Create a sequence of functions that uses aoi, start and end date to download the S2 data and add a cloud and shadow mask band - Python; colab.research.google.com; (jdbcode, 2022)
    - Request S2 and s2cloudless data products and merge to get an image collection of S2 SR with a cloud probability property
    - Using the cloud probability property and a cloud probability threshold, add a cloud mask band
    - Add bands for dark pixels, cloud projection, and cloud shadows
        - Use SCL band to identify water, Band 8 (NIR) and a NIR dark threshold to identify dark pixels, mean solar azimuth angle property and cloud projection distance threshold to identify cloud shadow projection distance and cloud shadows
    - Add a final band for the cloud and shadow mask
- First use data request and merge function to get image collection, then map the cloud and shadow mask functions to all images in image collection - Python; colab.research.google.com
- Reduce image collections to single images: take median of collection and clip to field's aoi - Python; colab.research.google.com
- Export images as .tif files from GEE (server side) to school google drive account (client side) - Python; colab.research.google.com
- Download images from google drive to local machine
- Import images as raster stacks into R (R Core Team, 2021)
- Calculate mean reflectance of the pixels value for each S2 band - R
    - Need to divide pixel values by 10000
- Calculate the five EOM indices - R
    - EOMI1 = B11 - B8A/B11 + B8A
    - EOMI2 =B12 - B4/B12 + B4
    - EOMI3 =(B11 - B8A) + (B12 - B4)/(B11 + B8A + B12 + B4)
    - EOMI4 =B11 - B4/B11 + B4
    - NBR2 = B11 - B12/B11 + B12
- Calculate separability index, i.e. euclidean distances between the pre and post index
    - I consulted with other grad students and we are not able to figure out an equation for the separability index. I emailed the author asking for this information but have not heard back.
- Perform paired two tailed T-test on the winning index - R
    - Paired two tailed test - Wilcox test (non-parametric) - http://www.sthda.com/english/wiki/paired-samples-wilcoxon-test-in-r
- Make figures and tables - R

Modeling Framework

The objective of the modeling framework is to *interpolate*, i.e. fill in gaps in knowledge of an environmental process (while the environmental process here is the manure application event, technically an anthropogenic process, by filling in gaps in the spatial and temporal occurrence of these events we hope to increase water quality modeling capabilities). Two methods were used for detecting changes, thus detecting these events, between the pre-and post application images. The first method is a non-statistical comparison of the fields' mean pixel value for pre and post application for each of the S2 spectral bands. While a paired T-test could be used to compare the pre and post means, a simple visual examination of the plot of the change in reflectance percentage for each band number is used.

The second method first calculates five spectral band indices originally developed for detecting SOC and crop residues. The histograms of each field's set of pre and post application pixel values for the five indices are examined, then the Euclidean distances between the histograms are calculated. The index with the largest distance between pre and post application is then chosen for a paired two tailed T-test. This is a basic statistical model to test for a significant difference in the pixel distributions for the index between pre and post application.

As it relates to machine learning model building, a T-test model does not require any training or validation. Simply, the test against a chosen significance level tells us if our methodology works. If we find a significant difference using this test, we might then be able to use the index as the predictand to train a machine learning workflow to detect manure applications from a time series of remote sensing data. For example, (Chugg et al., 2022) used hand labeled images to train a model to detect manure application events. However, if we could establish a spectral index that is highly correlated to manure application, we could use this to generate a much larger training data set for the same type of manure application event detecting model. Given the early stages of the use of multi-spectral imagery for detecting manure application, there is still work to be done to discover appropriate band indices that are correlated with the environmental process.

Results

Mean Reflectance Spectra Between Pre and Post Application Images

The difference in mean reflectance spectra between the pre and post application images was greatest at DCS in 2019 for the B6 band, AHS in 2020 for the B9 band, and at DCS in 2020 for the B8A band (Figure 1, Table 2). The smallest decreases were in bands B1 and B12. The smallest overall changes between pre and post occurred at DC in 2019, however this was the only field with data in this year.

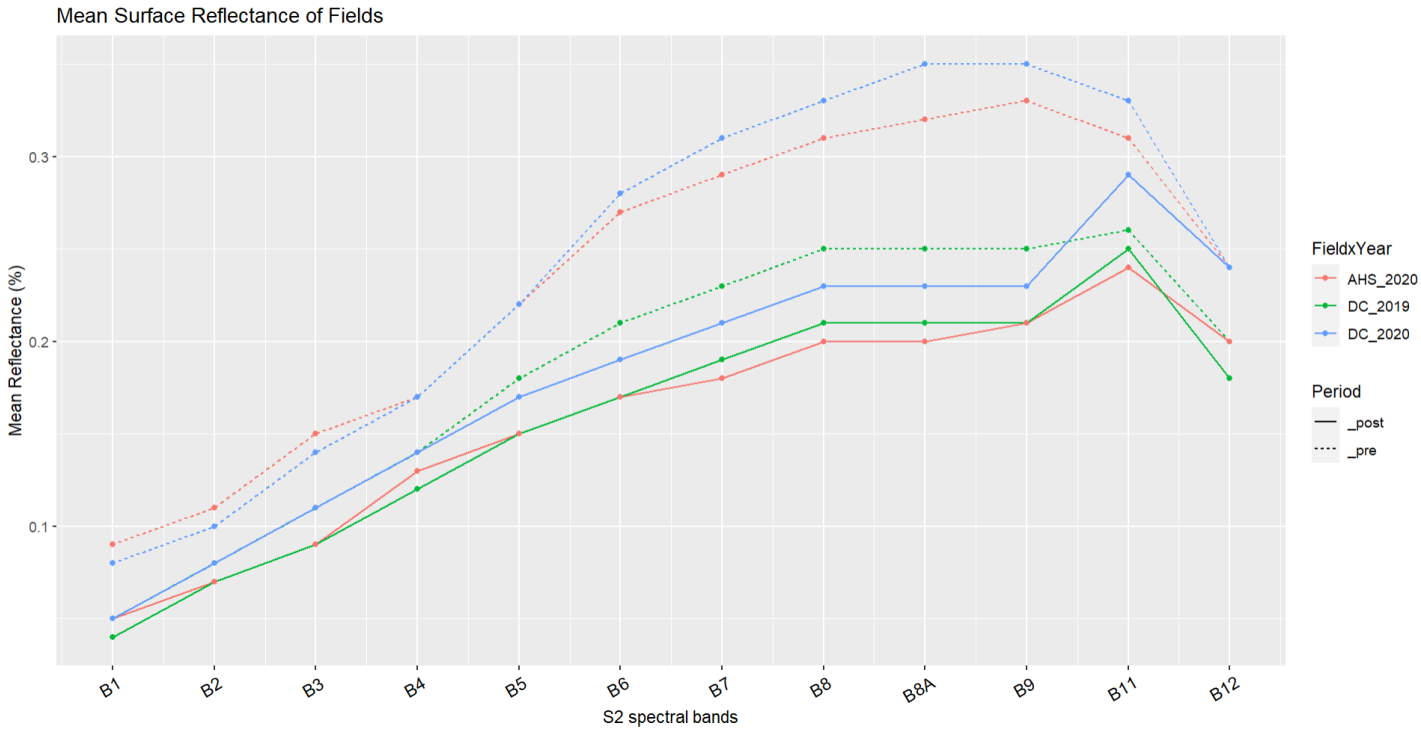

Figure 1 - Mean surface reflectance spectra of each field, year, and period by S2 band.

Table 2 - Minimum and maximum difference between pre and post reflectance spectra for each field + year combination.

	FieldxYear	min_diff	min_col	max_diff	max_col
1	AHS_2020	0.041	B1	0.123	B9
2	DC_2019	0.004	B1	0.039	B6
3	DC_2020	0.000	B12	0.114	B8A

Spectral Band Indices

Boxplots of the five spectral band indices are shown in Figure 2.The four EOM indices all showed increases in the median index value between pre and post application, while the NBR2 index showed a mix of increase and decrease across field + year combinations. A paired two tailed T-test was performed for all field + year combinations and all tests yielded significant differences between the pre and post indices (Table 3).

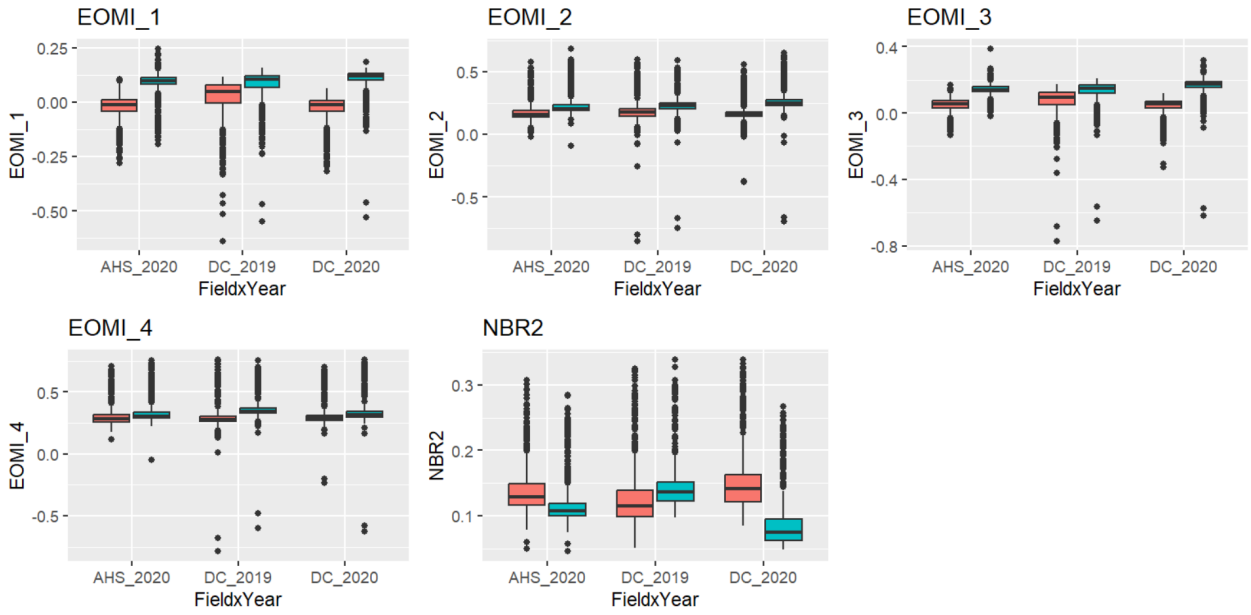

Figure 2 - Boxplots of band indices, red is the pre application image and blue is the post.

Table 3 - Wilcox test p-values for pre and post index values for each field + year combination

	FieldxYear	Index	Wilcox
1	AHS_2020	EOMI_1	0.000000e+00
2	AHS_2020	EOMI_2	7.140695e−290
3	AHS_2020	EOMI_3	0.000000e+00
4	AHS_2020	EOMI_4	1.615376e−155
5	AHS_2020	NBR2	9.051435e−262
6	DC_2019	EOMI_1	1.477443e−98
7	DC_2019	EOMI_2	1.071338e−60
8	DC_2019	EOMI_3	7.836195e−98
9	DC_2019	EOMI_4	3.752907e−72
10	DC_2019	NBR2	2.085296e−46
11	DC_2020	EOMI_1	1.237527e−96
12	DC_2020	EOMI_2	3.732227e−94
13	DC_2020	EOMI_3	1.360766e−96
14	DC_2020	EOMI_4	1.345776e−70
15	DC_2020	NBR2	3.983342e−99

Discussion

The shape of the curves in Figure 1 match those in (Dodin et al., 2021), however the mean reflectance values here are a bit lower. This could be due to the fact that in (Dodin et al., 2021) the manure applications were in the summer and here they were in the fall. Also, a noticeable difference in reflectance spectra for the three visible spectrum bands (Red:B4, Green:B3, Blue:B2) and NIR band (B8) across all field + years indicates that manure application at these sites could possibly be detected using 4 band imagery like NAIP. Thus, the supervised learning workflow proposed by (Chugg et al., 2022) may be expanded to our study location.

Boxplots here and in (Dodin et al., 2021) for EOMI 2 both showed increases in the index value between pre and post images. Significant differences between the pre and post indices was found here and by (Dodin et al., 2021). (Dodin et al., 2021) also had a control field (which received no manure applications for the past 30 years) to compare the post application indices with, and they also found significant difference between the two groups. A control field was not included here due to a lack of information on the management of neighboring fields.

The separation index was not able to be calculated here because the equation was not given in (Dodin et al., 2021), thus a 'winning' index was not able to be determined. Nevertheless, we showed that all indices had significantly different values between pre and post application, while (Dodin et al., 2021) only reported the differences for EOMI 2 since that was the best index as per their separation index calculation. The finding here, i.e. that all indices had significant differences between pre and post application, suggests that a wide range of indices could be used as predictors for training a multi-spectral detection model.

One of the major assumptions made here is that the median of the pre and post image collections was used to represent the field pixels. The medians were used because some images in the pre and post collections had pixels removed from the cloud and shadow masking. In (Dodin et al., 2021), only single images were used to represent the fields pre and post application, and this seems like the correct method.(Dodin et al., 2021) chose images that had no rainfall the day of acquisition, and reported minimal rainfall in the 3 and 7 days prior to the images. Rainfall and soil moisture are important factors to consider then using multi-spectral data to determine spectral shifts, which (Dodin et al., 2021) also showed in their field spectrometer experiment. A better workflow would have been to sort the image collections and use the ones with the fewest pixels removed by the mask.

The application events in this study were actually manure injections, i.e. six inches below the soil. This was assumed to not be an issue because it was observed during on-site visits the day after manure injection that manure also ponded on the surface. This was likely due to the fact that the heavy clay soils at these fields are prone to desiccation cracking and preferential flow pathways (PFPs) that start at the soil surface. Thus, despite the nosel injecting the pressurized manure slurry at depth, the pressurization also caused the manure to be forced up to the surface through the PFPs. On fields with more sandy soils, no manure is present on the surface after injection. Thus, only on fields with PFPs may this multi-spectral workflow be applicable to detect manure injection events, because multi-spectral imagery does not penetrate soil. In accordance with (Shea et al., 2022), the logical workflow for detecting manure injection events is to use radar data since it can penetrate the soil/quantify changes in soil moisture.

Conclusions

Understanding the spatial and temporal nature of manure applications is an important part of modeling nutrient loss and subsequent water quality. With the expectation of the rare circumstance where expert knowledge is available, dates and locations of application events are seldom known. Remote sensing has been used to detect changes in environmental processes, as well as quantify the ecological impacts of agriculture, thus it is a promising tool for filling gaps in knowledge regarding manure application locations and timing. Nevertheless, limited work has been done applying remote sensing techniques to detect manure applications. (Dodin et al.,

2021) showed that the reflectance spectra of fields was noticeably different after manure applications. The authors also applied spectral band indices to the images to discern an index that may be useful for detecting changes between images and thus manure applications.

The methods of (Dodin et al., 2021) were applied to a dairy farm in Vermont, USA, to test the replicability and perform a novel analysis, i.e. has not been conducted in this region. Through replicating their analysis, it was discovered that their results were unable to be fully replicated due to insufficient documentation of their methods. Apart from that, the results here align with (Dodin et al., 2021) in that multi-spectral imagery has potential to detect manure application events. More work needs to be done to determine a winning band index. As an aside, from the findings in this project, it is possible that a different index from the one chosen in (Dodin et al., 2021) is ideal for identifying spectral shifts. More thought should be given to identify a winning index given there is limited work on the subject.

References

Chugg, B., Rothbacher, N., Feng, A., Long, X., & Ho, D. E. (2022). *Detecting Environmental Violations with Satellite Imagery in Near Real Time: Land Application under the CleanWater Act* (Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22)) [Arxiv]. https://doi.org/10.1145/3511808.3557104

Dodin, M., Smith, H. D., Levavasseur, F., Hadjar, D., Houot, S., & Vaudour, E. (2021). Potential of Sentinel-2 Satellite Images for Monitoring Green Waste Compost and Manure Amendments in Temperate Cropland. *Remote Sensing*, *13*(9), 1616. https://doi.org/10.3390/rs13091616

Easton, Z. M., Kleinman, P. J. A., Buda, A. R., Goering, D., Emberston, N., Reed, S., Drohan, P. J., Walter, M. T., Guinan, P., Lory, J. A., Sommerlot, A. R., & Sharpley, A. (2017). Short-term Forecasting Tools for Agricultural Nutrient Management. *Journal of Environmental Quality*, *46*(6), 1257–1269. https://doi.org/10.2134/jeq2016.09.0377

Fu, B., Horsburgh, J. S., Jakeman, A. J., Gualtieri, C., Arnold, T., Marshall, L., Green, T. R., Quinn, N. W. T., Volk, M., Hunt, R. J., Vezzaro, L., Croke, B. F. W., Jakeman, J. D., Snow, V., & Rashleigh, B. (2020). Modeling Water Quality in Watersheds: From Here to

the Next Generation. *Water Resources Research*, *56*(11). https://doi.org/10.1029/2020WR027721

Haith, D. A., & Atkinson, D. W. (1977). A linear programming model for dairy farm nutrient management. *Journal of Environmental Quality*, *7*(4).

jdbcode. (2022, December 16). *Semtinel-2 Cloud Masking with s2cloudless* [Tutirial]. Earth Engine Developer Community. https://developers.google.com/earth-engine/tutorials/community/sentinel-2-s2cloudless

R Core Team. (2021). *R: A language and environment for statistical computing* [R Foundation for Statistical Computing]. https://www.R-project.org/

Shea, K., Schaffer-Smith, D., & Muenich, R. L. (2022). Using remote sensing to identify liquid manure applications in eastern North Carolina. *Journal of Environmental Management*, *317*, 115334. https://doi.org/10.1016/j.jenvman.2022.115334