Counterfactual Analysis of Neural Networks Used to Create Fertilizer Management Zones

Giorgio Morales and John Sheppard
Gianforte School of Computing
Montana State University
Bozeman, MT 59717

***Abstract*—In Precision Agriculture, the utilization of management zones (MZs) that take into account within-field variability facilitates effective fertilizer management. This approach enables the optimization of nitrogen (N) rates to maximize crop yield production and enhance agronomic use efficiency. However, existing works often neglect the consideration of responsivity to fertilizer as a factor influencing MZ determination. In response to this gap, we present a MZ clustering method based on fertilizer responsivity. We build upon the statement that the responsivity of a given site to the fertilizer rate is described by the shape of its corresponding N fertilizer-yield response (N-response) curve. Thus, we generate N-response curves for all sites within the field using a convolutional neural network (CNN). The shape of the approximated N-response curves is then characterized using functional principal component analysis. Subsequently, a counterfactual explanation (CFE) method is applied to discern the impact of various variables on MZ membership. The genetic algorithm-based CFE solves a multi-objective optimization problem and aims to identify the minimum combination of features needed to alter a site's cluster assignment. Results from two yield prediction datasets indicate that the features with the greatest influence on MZ membership are associated with terrain characteristics that either facilitate or impede fertilizer runoff, such as terrain slope or topographic aspect.**

***Index Terms*—Neural network response curves, management zones, counterfactual explanations, explainable machine learning**

I. Introduction

In precision agriculture (PA), management zones (MZs) are distinct sub-regions in a field with similar yield-influencing factors [1]. Different MZs account for the variability of factors within the field (e.g., soil composition) and, thus, vary in their requirements for specific treatments. These zones are areas with relative homogeneity where specific crop management practices are implemented, aiming to optimize crop productivity and reduce the environmental impact by reducing the overall fertilizer applied [2], [3].

Several methods have been proposed for the delineation of MZs. Some of them rely on historical yield data solely [4], [5] while others use information extracted from remote sensing data exclusively [6], [7]. Alternatively, certain approaches employ a combination of covariate factors, encompassing various soil properties, environmental factors, and topographic information [8], [9], [10], [11]. Most of these methods are based on unsupervised learning techniques; specifically, clustering methods such as k-means [12] and fuzzy c-means [8], [9], [10], [13], and principal component analysis (PCA) [8], [9], [11]. In addition, MZ delineation methods based on supervised learning such as random forests (RFs) and support vector machines (SVMs) [7], [11] have emerged recently.

All previous works produce management zones using factors that are directly or indirectly related to crop productivity. Nevertheless, to the best of our knowledge, no published work has explicitly considered fertilizer responsivity as the main driver for defining MZs. One of the key objectives of using MZs is to equip farmers with the necessary tools to make informed decisions about crop management, such as determining the appropriate amount of fertilizer required in each zone. As a consequence, our efforts should be directed toward establishing management zones where all included sites display comparable responsivity to varying fertilizer rates [4].

Fertilizer responsivity can be characterized using nitrogen fertilizer-yield response (N-response) curves. These curves exhibit the estimated crop yield values corresponding to a specific field site in response to all admissible fertilizer rates, typically ranging between 0 and 150 pounds per acre (lbs/acre) [14], [15]. The shape of the N-response curve indicates the site's responsiveness to fertilizer, with a flat curve suggesting low responsivity and a steep curve suggesting high responsivity. Thus, in this work, we propose an MZ clustering method that accounts for within-field variability of fertilizer responsivity based on approximated N-response curves.

To do this, we derive non-parametric response curves from observed data. Most response curves are generated based on parametric assumptions using methods such as linear regression or quadratic plateau regression; however, our experience suggests that N-response is much more complex than these models can capture. Our approach is based on a convolutional neural network (CNN) acting as a regression model to map the covariate factors to crop yield, as suggested in [16]. Then, the CNN is used to generate approximated N-response curves across a range of admissible fertilizer rates [15]. The distinction in shape between two N-response curves is quantified by measuring the distance between the corresponding transformed curves in a reduced space, calculated using functional principal component analysis (fPCA). Thus, determining MZs relies on leveraging the shape dissimilarity of N-response curves as the key distance metric in cluster analysis.

It is worth pointing out that none of the existing methods for determining MZs attempt to provide explanations regarding the behavior of their results. This presents a significant limitation, particularly within the context of the growing

area of explainable artificial intelligence (XAI). Therefore, an approach to determining MZs that is inherently explainable should enable farmers to discern cause-and-effect relationships between their inputs and outputs, enabling a more transparent decision-making process. Therefore, we use a *post-hoc* explainability method to facilitate understanding the impact of various covariates on determining the MZ assignment for a given site. Specifically, we employ a counterfactual explanation (CFE) method, adapted from a previous work [15].

For a specific site within the field, the aim of CFE is to identify the minimal set of covariate factors that, when modified, leads to the response curve of the generated counterfactual sample being assigned to a cluster different from the original one. The problem of generating such CFEs is tackled as a multi-objective optimization problem (MOO) and is solved using an approach based on the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [17]. Applying this process across all sites associated with a specific management zone enables the computation of global relevance scores. These scores facilitate identifying covariates with the most significant impact on cluster membership assignment.

Our specific contributions are summarized as follows. First, we present the first MZ generation approach that accounts for within-field variability of fertilizer responsivity based on neural network-generated N-response curves derived from observed data. Second, we show how to apply a *post-hoc* explainability method that generates CFEs to reveal the influence of covariate factors on MZ assignments.

II. Related Work

Several studies in PA have tackled the challenge of determining MZs, employing a spectrum of techniques ranging from traditional statistical methods to machine learning algorithms. For instance, Georgi *et al.* [6] developed a cost-effective segmentation algorithm using multi-spectral satellite data. They generate normalized difference vegetation index (NDVI) maps and apply morphological operations to identify homogeneous regions. The NDVI values found within the processed maps are then classified into four quantiles, which are used to delineate four MZs. Two assumptions are made: a direct correlation exists between NDVI and crop yield, and MZs can be determined based on expected productivity.

Gallardo-Romero *et al.* [7] presented work based on similar assumptions, using vegetation indices extracted from multi-spectral images, along with available yield maps from previous growing seasons. The study categorizes yield values into three classes, assuming that each category corresponds to a distinct MZ. To predict productivity levels (i.e., MZs), the authors employ an ensemble of four machine learning algorithms: classification and regression trees (CART), RF, gradient boosting trees (GBT), and SVM.

To motivate our method, we argue that the similarity in estimated yield values for different sites under specific conditions does not imply that these sites would exhibit similar responses under all admissible conditions. For instance, take two sites whose predicted yield values are the same when using the same treatment and, as such, are classified within the same management zone. Assume we know their N-response curves, as depicted in Fig. 1. Despite both zones producing the same yield when using $N1$ lbs/ac, notice that they react to differing amounts of nitrogen differently. In fact, the second zone can achieve the same yield with less nitrogen ($N2$)!

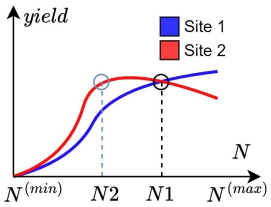


Fig. 1: N-response curve comparison for two sites with the same predicted yield value given a treatment of $N1$ lbs/ac.

Instead of focusing on predicting productivity levels, other methods have presented techniques to cluster homogeneous regions based on the characteristics of the field. Fridgen *et al.* [13] introduced the Management Zone Analyst (MZA), a tool that utilizes a fuzzy classification procedure. MZA also provides performance indices such as the fuzziness performance index (FPI) and normalized classification entropy (NCE) to aid in determining the optimal number of clusters for MZs. This tool has been applied with other clustering methods for determining MZs based on soil traits, NDVI data, and multi-year crop yields [10].

Fuzzy c-means clustering, often paired with PCA, has been used widely for determining site-specific MZs [8], [9], [18]. These methods take into account the multi-dimensional nature of soil variability, allowing for more meaningful zone delineation. Other approaches involve integrating fuzzy c-means and PCA with other machine learning algorithms, such as RF. For example, Maleki *et al.* [11] trained an RF model to predict various soil properties (e.g., pH, sand content, and electro-conductivity) given environmental covariates (e.g., vegetation indices, geology maps, and water quality parameters). The RF model was used for feature selection to remove unimportant environmental covariates. PCA was applied to the reduced features, and the MZs were identified using fuzzy c-means.

Other previous methods overlooked the importance of the explainability aspect concerning the resulting MZs. Addressing this gap, our work tackles the issue by incorporating counterfactual explanation analysis, aiming to uncover causal dependencies between inputs and outputs. Nevertheless, there has been limited focus on CFE methods in the context of regression tasks. Classification-based CFE methods aim to produce counterfactual samples with predicted class labels that differ from the original ones [19]. In this context, identifying features of higher relevance involves recognizing those undergoing more frequent changes during CFE generation [20].

Based on this concept, we proposed a CFE method [15] that, given a response curve generated for a selected input feature (known as the "active feature"), allows for the identification of the remaining system variables (known as the "passive

features") with the highest relevance on the curve's shape. This method diverges from classic sensitivity analysis or saliency analysis, typically employed to explore how various values of a set of independent variables influence the response variable. In contrast, our approach delves into understanding how different combinations of the passive features influence the response variable across the entire range of admissible values of the active feature, without assuming independence among the features. In this work, we adapt this method to understand the impact of passive features on the MZ membership of a field site. Thus, a new objective function will be presented.

III. Materials and Methods

A. Datasets

We utilized two datasets for early-yield prediction acquired from two winter wheat dryland fields called "Field A" and "Field B". These datasets were compiled and discussed in a previous work [16]. The problem of predicting crop yield is formulated as a regression task, with a set of eight covariates collected in March serving as the explanatory variables:

1) N: Nitrogen rate (lb/ac).
2) S: Topographic slope (degrees).
3) E: Topographic elevation (m).
4) TPI: Topographic position index.
5) A: Topographic aspect (radians).
6) P: Precipitation from the prior year (mm).
7) VV and VH: Backscattering coefficients derived from synthetic aperture radar (SAR) images from Sentinel-I.

The dependent variable is the harvested yield in bushels per acre (bu/ac), recorded post-harvest in August. Therefore, the March-acquired data serve as predictors for the yield outcomes obtained in August. The acquired feature and yield maps can be seen as image rasters, where each pixel or cell represents a region of 10×10 m. Data were collected across three growing seasons for each field (2016, 2018, and 2020).

B. Convolutional Neural Network

In [16], we tackled the yield prediction problem using a two-dimensional (2D) regression model; that is, a model with 2D inputs and 2D outputs. We designed a 3D–2D CNN architecture called Hyper3DNetReg, which is trained to predict the yield values of all cells within a small spatial neighborhood of a field simultaneously. Specifically, our model takes as input an image data cube of 5×5 pixels with $n = 8$ channels, where each channel represents a distinct covariate. The model then generates a 2D image output of 5×5 pixels.

Hence, data from each field were partitioned into several 5×5–pixel patches to create their corresponding training and validation sets. The training dataset comprised 90% of the patches extracted from three available years of data, with the remaining 10% constituting the validation dataset. Our trained models are field-specific, signifying that they are trained on data from a particular field and employed to predict future yield maps using data from the same crop and the same field.

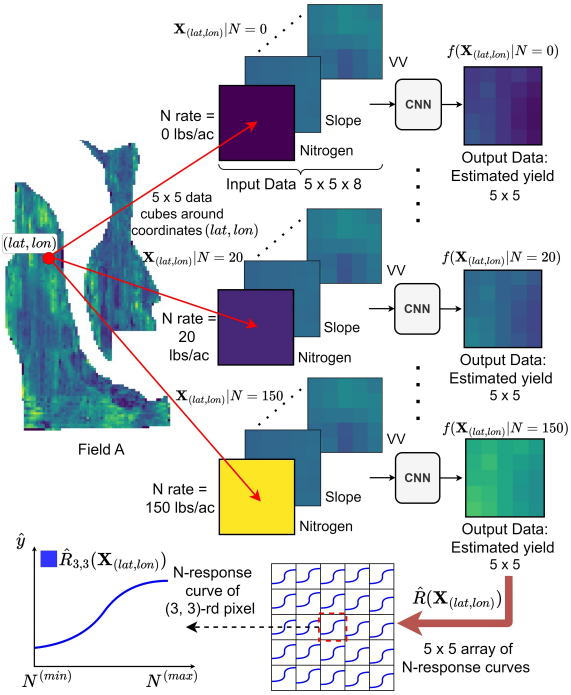


Fig. 2: Generation of a 5×5 array of N-response curves generated around a field point at coordinates (lat, lon).

C. N-response Curve Generation

After the network has been trained and, assuming it captures the underlying causal structure of the problem adequately, it can be employed to produce approximate response curves. In previous work [15], we defined a response curve as a tool that allows for the responsivity analysis of a sensitive system to a particular "active feature". In addition, other stimuli that may influence the relationship between the response variable and the active feature were referred to as "passive features". In the context of the generation of N-response curves, the active feature corresponds to N (nitrogen rate), while the seven remaining variables represent the set of passive features.

An input data cube is represented as $\mathbf{X} = \{X_1, \ldots, X_n\}$, with X_1 corresponding to the N covariate, and the subsequent dimensions aligning with the remaining covariates based on the order outlined in Section III-A. Let us denote the trained CNN model as $f(\cdot)$ so that, given an input $\mathbf{X}$, its estimated yield patch is denoted as $\hat{\mathbf{Y}} = f(\mathbf{X})$. Consider we produce estimated yield patches for all admissible values of N (bounded by $N^{\min} \leq N \leq N^{\max}$) and stack them as a data cube $\hat{R}(\mathbf{X})$:

$$\hat{R}(\mathbf{X}) = \{f(\mathbf{X}|N = N^{\min}), \ldots, f(\mathbf{X}|N = N^{\max})\}, \quad (1)$$

as shown in Fig. 2. As such, the (i, j)-th cell of $\hat{R}(\mathbf{X})$ (where $1 \leq i, j \leq 5$), denoted as $\hat{R}_{i,j}(\mathbf{X})$, represents the approximated response curve corresponding to the (i, j)-th cell of input $\mathbf{X}$.

The example in Fig. 2 shows the response curve generation process of all pixels within input $\mathbf{X}_{(lat,lon)}$, which represents the 5×5–pixel patch generated around coordinates (lat, lon) of the field. However, our goal is to generate N-response curves for all sites within the field. To do so, we move

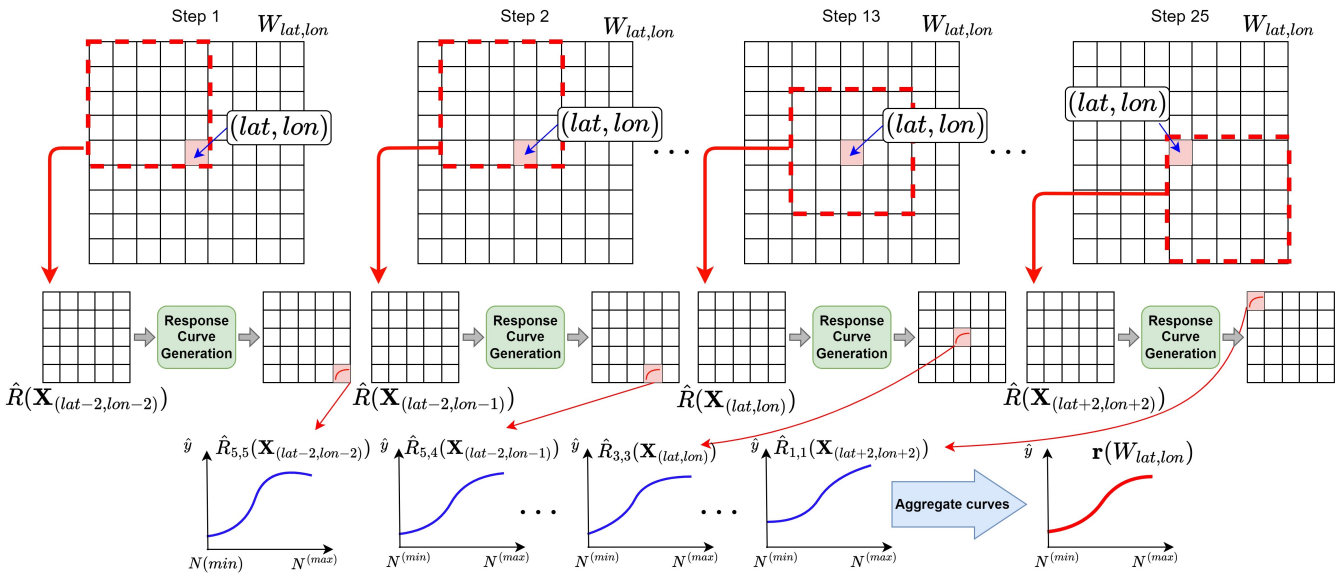


Fig. 3: N-response curves aggregation for a field point at coordinates (lat, lon).

a 5×5–pixel sliding window throughout the entire field. Note that this approach results in overlapping predicted yield patches for consecutive points. Therefore, results obtained from neighboring points of the field must be aggregated.

Fig. 3 depicts the aggregation process of N-response curves obtained for a field point at coordinates (lat, lon). This process involves considering all valid neighboring 5×5–pixel patches; i.e., patches whose centers are located within the field and that contain the point at (lat, lon), highlighted in red. The 9×9 window generated around the field point at (lat, lon) is denoted as $W_{lat,lon}$. For each of the valid patches in $W_{lat,lon}$, a 5×5 array of N-response curves is generated using Eq. 1. Then, the N-response curves corresponding to the field point at (lat, lon) are averaged, yielding a singular approximated N-response $\mathbf{r}(W_{lat,lon})$. This averaging process alleviates noisy outcomes and produces smoothed curves.

In Sec. I, we stated that the fertilizer responsivity of a given site is characterized by the shape of its N-response curve. As such, when comparing the shape of two or more N-response curves, the focus is not on their absolute estimated yield values. Hence, any vertical shifts are eliminated to obtain the aligned approximate N-response curve $\tilde{\mathbf{r}}(W_{lat,lon})$ as follows:

$$\tilde{\mathbf{r}}(W_{lat,lon}) = \mathbf{r}(W_{lat,lon}) - \min(\mathbf{r}(W_{lat,lon})).$$

D. Functional Principal Component Analysis

We compute the set $\mathbf{R}$ comprising the aligned approximate N-response curves generated for all sites within the field. $\mathbf{R}$ constitutes a set of functional data whose samples are approximated N-response curves. Thus, fPCA can be applied to $\mathbf{R}$ to establish a distance metric conveying the difference in shape between N-response curves, as suggested in [15].

Functional Principal Component Analysis extends traditional PCA to analyze and represent variability in functional data [21]. As such, an N-response curve can be expressed as a linear combination of *functional* principal component (fPCs). Each fPC encapsulates a unique curve pattern, implying that curves with distinct shapes will be encoded using different fPC values. In this work, we suggest approximating an N-response curve using $K = 3$ fPCs, a choice justified by their ability to explain at least 99.5% of the variance of fields A and B. Thus, the proposed distance metric between curves $\mathbf{r}_1$ and $\mathbf{r}_2$ is:

$$d(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{\sum_{k=1}^{K} (v_k(\mathbf{r}_1) - v_k(\mathbf{r}_2))^2}, \tag{2}$$

where $v_k(\mathbf{r}_j)$ is the value of the k-th principal component obtained after transforming the curve $\mathbf{r}_j$.

E. Management Zone Clustering

Using fuzzy c-means has become a prevalent approach in management zone delineation methods [8], [9], [18]. In fuzzy c-means, each data point is assigned a membership score indicating the extent to which it belongs to a specific cluster. A cluster centroid is computed as the mean of all data points, weighted by their respective cluster membership values.

We propose to cluster all field points based on their fertilizer responsivity so that each cluster corresponds to a distinct MZ. The process involves generating aligned approximate N-response curves for all field sites, followed by their transformation into a reduced three-dimensional (3D) space through fPCA. Hence, the difference in fertilizer responsivity between curves (i.e., the difference in curve shape) is conveyed by their Euclidean distance in the transformed space. Therefore, the fertilizer responsivity distance (Eq. 2) serves as the distance metric for the fuzzy c-means algorithm.

Some approaches utilize indices such as the silhouette score, fuzziness performance index, and normalized classification entropy to determine the optimal number of clusters [10], [13]. However, these indices might face challenges in situations where clusters lack clear separation, as observed in the present

context. Recall that all data points for clustering belong to the same field, leading to gradual changes in soil variability and, as a consequence, gradual changes in fertilizer responsivity.

In addition, in PA, it is a common practice to specify between three and five MZs [7]. The decision to use up to five MZs is often influenced by practical considerations, such as the limitations of variable rate application machinery and the complexity of the field. For instance, using more than five zones may entail intricate zone boundaries, posing challenges for certain variable rate technologies to distinguish between closely situated zones. Following this convention, we chose through visual inspection a cluster count that minimizes the creation of redundant or highly variable MZs. In future work, we will design statistical tests that will determine if the responsivity of multiple clusters of functional data is significantly different.

F. Counterfactual Explanations for Management Zones

Consider a field point at coordinates (lat, lon); the 9×9–pixel window generated around it, $W_{lat,lon}$; and the aligned N-response curve generated for the center of this window, $\tilde{\mathbf{r}}(W_{lat,lon})$. Let $W'_{lat,lon}$ represents a counterfactual explanation of $W_{lat,lon}$ and let $\tilde{\mathbf{r}}(W'_{lat,lon})$ be its corresponding aligned N-response curve. The CFE $W'_{lat,lon}$ is generated by introducing perturbations to the original set of passive features in $W_{lat,lon}$, thereby altering the cluster membership of the resulting N-response curve $\tilde{\mathbf{r}}(W'_{lat,lon})$. Furthermore, we seek to identify the minimal set of passive features requiring perturbation for a shift in cluster membership, which focuses on identifying features with the highest impact on this alteration.

This problem is tackled as an MOO problem. To solve it, we adapt the method presented in [15]. Originally, this method was proposed to minimize three competing objectives using NSGA-II [17]. Its primary use is to identify the minimal set of passive features requiring modification, ensuring that the distance between the responsivity of the counterfactual explanation (CFE) and that of the original samples exceeds a predefined hyperparameter threshold. Note that this approach does not account for management zone membership and is used to analyze the overall field behavior.

Specifically, our MOO problem is defined as:

$$\min_{W'_{lat,lon}} \; (g_1(W_{lat,lon}, W'_{lat,lon}), g_2(W_{lat,lon}, W'_{lat,lon}), g_3(W_{lat,lon}, W'_{lat,lon})).$$

The first objective differs from the first objective presented in [15], and aims for a shift in cluster membership, such that:

$$g_1(W, W') = \begin{cases} -1, & \text{if } \texttt{cl}(\tilde{\mathbf{r}}(W')) \neq \texttt{cl}(\tilde{\mathbf{r}}(W)) \wedge \\ & \quad \texttt{m}(\tilde{\mathbf{r}}(W')) > \epsilon \\ 0, & \text{otherwise,} \end{cases}$$

where $\texttt{cl}(\tilde{\mathbf{r}}(W'))$ is the cluster assignment for a curve $\tilde{\mathbf{r}}(W')$ after undergoing transformation via fPCA, while $\texttt{m}(\tilde{\mathbf{r}}(W')) > \epsilon$ denotes the membership score that $\tilde{\mathbf{r}}(W')$ belongs to the assigned cluster. Recall that a field point is assigned a membership score for each possible cluster by fuzzy c-means.

Notice that the objective is to alter the cluster membership so that the membership score assigned to the new cluster is greater than ϵ, which is a tunable hyperparameter. The reasoning behind the decision is as follows. Suppose a curve $\tilde{\mathbf{r}}(W)$ is assigned to a certain cluster with low membership; then, this curve would be susceptible to changes in cluster membership when subjected to minor random perturbations. As a result, such changes may not be informative or representative of the behavior of the remaining curves within the same cluster. To avoid this situation, we enforce that $\texttt{m}(\tilde{\mathbf{r}}(W')) > \epsilon$ (e.g. $\epsilon > 0.8$) during CFE generation. This ensures that the CFE is assigned to a different cluster with high confidence, making the necessary changes more meaningful.

The second objective is related to minimizing the number of modified features:

$$g_2(W, W') = ||W - W'||_0,$$

where $||\cdot||_0$ represents the L_0 norm. Finally, the third objective accounts minimizing the distance between the original sample and its CFE. Considering the real-valued nature of the covariates in this study, each exhibiting distinct ranges of values, the distance is calculated as:

$$g_3(W, W') = \frac{1}{n} \sum_{s=1}^{n} \frac{1}{r_s} |W^{(s)} - W'^{(s)}|,$$

where $W^{(s)}$ is the s-th covariate, and r_s its range of values.

This MOO problem is solved using NSGA-II with a population size of T_0 CFE candidate solutions. NSGA-II outputs a set of T Pareto-optimal solutions ($T \leq T_0$). These solutions are non-dominated, meaning that no other solution in the set is superior in all objectives simultaneously. Hence, our selection criteria for a solution from the Pareto set involve choosing solutions with the minimum g_1 value, followed by those with the lowest g_2 value, and finally selecting solutions with the minimum g_3 value.

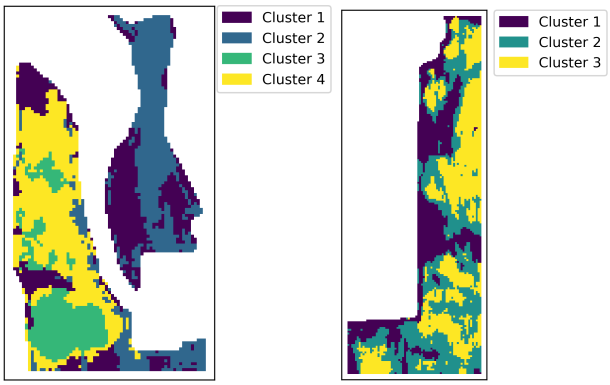
Furthermore, we calculate local and global explanations for the generated results. Given a field point, a local explanation identifies the passive features with the greatest impact on its management zone membership. As such, a local explanation for a field point at coordinates (lat, lon), denoted as $\alpha_{lat,lon}$ is given by the set of features that were altered during the generation of the counterfactual sample:

$$\alpha_{lat,lon} = \{s \mid W^{(s)}_{lat,lon} \neq W'^{(s)}_{lat,lon}\}.$$

In contrast, global explanations are generated for *each MZ* and convey the relevance of passive features at the MZ level. When analyzing the z-th MZ, the individual feature relevance of the s-th passive feature is calculated as follows:

$$r_z^{(s)} = \frac{1}{|\mathbf{C}_z|} \sum_{(lat,long) \in \mathbf{C}_z}^{|\mathbf{C}_z|} \mathbb{I}_{s \in \alpha_{lat,lon}},$$

where $\mathbf{C}_z$ represents the set of field points that have been clustered into the z-th MZ. Thus, $r_z^{(s)}$ is the ratio of times that the s-th passive feature was modified when generating CFEs for field points in $\mathbf{C}_z$.



(a) (b)

Fig. 4: Management zones for (a) Field A and (b) Field B.

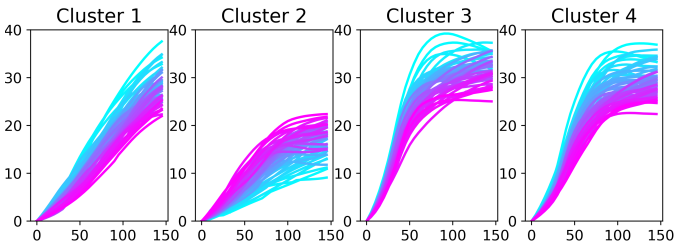


Fig. 5: Aligned approximated N-response curves corresponding to each management zone generated for Field A.

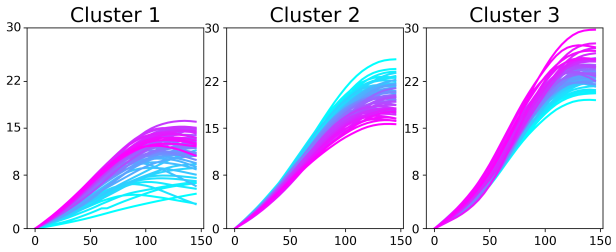


Fig. 6: Aligned approximated N-response curves corresponding to each management zone generated for Field B.

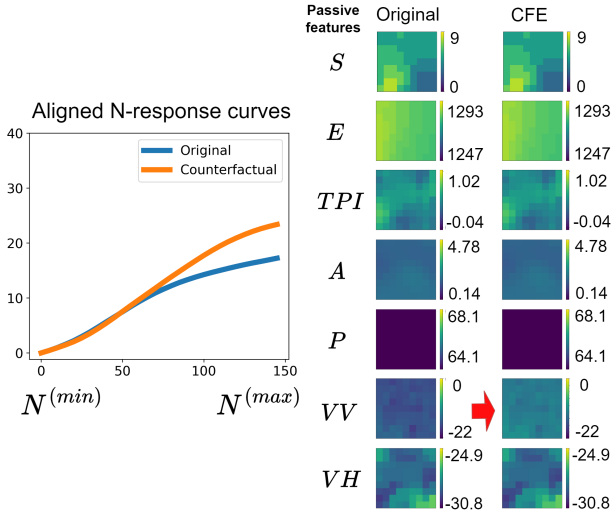


Fig. 7: Example of a counterfactual N-response curve generated for a field point of management zone "2" of Field A.

Finally, we emphasize that, in this work, we did not assume mutual independence among the covariate variables. During the generation of CFEs, more than one passive feature may be modified simultaneously. This suggests that certain combinations of features are more effective than modifying individual features in isolation. Therefore, to analyze which features react together and identify the most effective combinations, we report the five most frequently occurring feature combinations.

IV. Experimental Results

We evaluated our MZ clustering method outlined in Sec III-E on Fields A and B. For Field A, we decided to generate four MZs (i.e., four clusters). Conversely, for Field B, we chose to generate three MZs. This decision is justified by the fact that Field B is more homogeneous than Field A, thus their corresponding N-response curves show less variability. The MZs obtained for Fields A and B are shown in Figs.4a and 4b, respectively. In addition, Fig.5 and Fig.6 show fifty aligned approximate response curves selected randomly from each MZ obtained from each field. The implementation code is available at https://github.com/NISL-MSU/ManagementZonesCFE.

Furthermore, we applied our CFE analysis method to all field points of all MZs created for Fields A and B. For NSGA-II, we used a population size of $T_0 = 50$ and 100 iterations. For objective $g_1(\cdot)$, we selected $\epsilon = 0.8$. For example, Fig. 7 illustrates a counterfactual N-response curve generated for a site of Field A that was initially clustered into MZ "2." This CFE exhibits an increase in the VV variable, resulting in an N-response curve with higher responsivity and a shift in cluster membership to "1." We conducted experiments with alternative ϵ values, such as 0.85 and 0.9, and observed similar results, although they led to slower convergence rates. These findings are not included in this paper due to space limitations.

Finally, we show the calculated individual feature relevance values obtained for all passive features of Fields A and B in Fig. 8 and Fig. 9, respectively. In addition, TablesI and II list the five most repeated feature combinations found during the CFE analysis of each MZ.

V. Discussion

Our method involves using CNNs for the automated generation of approximated N-response curves, which are employed to describe the fertilizer responsivity at different locations in the field. Fig. 5 and Fig. 6 show a subset of these non-parametric curves, demonstrating a behavior aligned with agronomic expectations; i.e., sigmoid-*like* curves that capture an apparent yield loss after reaching a saturation point.

These figures confirm that our MZ determination method organized the generated N-response curves effectively into clusters with consistent curve shapes within each cluster. The observed variation in curve shapes across different clusters highlights different patterns of fertilizer responsivity. We claim that identifying MZs characterized by specific fertilizer

TABLE I: Top-five feature combinations – Field A

Zone	1		2		3		4	
#	Comb.	% Rep.	Comb.	% Rep.	Comb.	% Rep.	Comb.	% Rep.
1	[S]	39.5	[VV]	50	[TPI]	53	[S]	41
2	[A]	22	[S]	16	[S]	33	[A]	23.5
3	[VV]	10.5	[A]	8	[VV]	5.5	[VV]	16.5
4	[S, A]	5	[TPI]	7	[A]	5.5	[TPI]	13
5	[A, VV]	3.5	[VH]	4	[S, VV]	1.5	[S, A]	2

TABLE II: Top-five feature combinations – Field B

Zone	1		2		3	
#	Comb.	% Rep.	Comb.	% Rep.	Comb.	% Rep.
1	[S]	66.5	[S]	49	[S]	55.5
2	[S, VV]	11	[VV]	23	[TPI]	26.5
3	[TPI]	6	[TPI]	13.5	[VV]	9
4	[S, A]	3	[S, VV]	4.5	[S, TPI]	3
5	[S, TPI]	1.5	[A]	3	[VV, TPI]	1.5

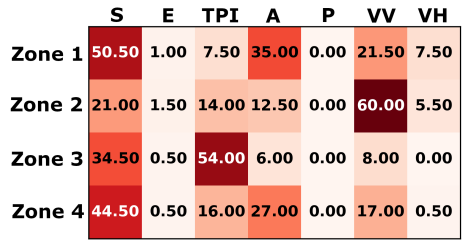


Fig. 8: Individual feature relevance values of Field A.

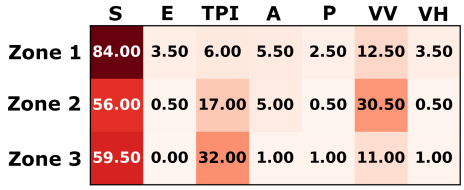


Fig. 9: Individual feature relevance values of Field B.

responsivity patterns offers valuable insights for designing treatments tailored to the specific needs of each MZ.

In Fig. 5, we observe that cluster "2" represents an MZ encompassing areas of the field characterized by low fertilizer responsivity. These areas exhibit a limited reaction to changes in fertilizer rate compared to the other zones. In contrast, cluster "3" represents the MZ exhibiting the highest fertilizer responsivity. Note that, while the curves in cluster "4" share a similar shape with those in cluster "3", the latter depicts a slightly steeper behavior, achieving higher yield values with lower fertilizer rates. Similarly, the curves depicted in Fig. 6 can be associated with low, medium, and high fertilizer responsivity regions. In our initial experiments, we found that introducing a fourth MZ would lead to the creation of a redundant cluster with curves closely resembling those in cluster "3". We observed that employing four MZs resulted in zone boundaries that were overly intricate and impractical.

Note that the adapted CFE generation method enables us to derive local explanations for each field location, providing insights into why they were assigned to a specific MZ. For instance, Fig. 7 shows a site of Field A that was assigned to cluster "2". The generated CFE suggests that an increase in the VV variable could enhance the site's fertilizer responsivity, leading to its clustering into another MZ. Given that VV is associated with soil moisture content, this outcome could be interpreted as if the primary factor contributing to the site's assignment to cluster "2" is its relatively low moisture content.

Applying the CFE generation process to all field points and aggregating the results enables us to derive global explanations that characterize the overall behavior of each MZ. For example, from Fig. 8 and Table I, we conclude that variables S, A, and VV have the most significant influence in determining the assignment of a particular field point to MZ "1." Fig. 10 shows the topographic slope and the topographic aspect maps for Field A. From this, we observe that sites clustered in MZ "1" coincide with areas with high aspect values and medium to high slope values, as indicated by the red ellipses. Note that areas with higher slope values are more prone to fertilizer runoff, which affects their fertilizer responsivity. In addition, high aspect values correspond to a northeast slope orientation (in the Northern Hemisphere), implying limited sunlight and increased snow retention in comparison to other regions of the field. From Fig. 8 and Table I, we also conclude that the most relevant variables in determining the assignment of a particular field point to MZ "3" (i.e., the highest fertilizer responsivity) are TPI and S. Note in Fig. 10 that the sites in MZ "3" occupy areas with low slope values, as well as most of the sites in MZ "4". These findings indicate that TPI plays a crucial role in distinguishing the fertilizer responsivity of relatively flat regions. TPI is associated with the terrain's ruggedness, and higher values correspond to an increased likelihood of runoff. Similar explanations are obtained for the rest of the MZs.

On the other hand, the results from Fig. 9 indicate that the feature relevance values are similar for the three MZs. This can be attributed to the uniform elevation values observed throughout Field B, where there are no steep and abrupt regions such as in Field A. Therefore, similar to the case of the MZ "3" of Field A, the most important variables for cluster membership are S, TPI, and VV, according to Fig. 9 and Table II. It is worth pointing out that, for both fields, most of the generated CFEs required changes in individual features, as evidenced in Table I and Table I. However, in the case of Field B, there was a greater need for changes involving at least two variables compared to Field A. For instance, the combination of S and VV was the second most effective variable combination, recurring in 11% of the CFEs generated for MZ "1." This implies that, in many cases, the low responsivity of sites in MZ "1" can be attributed not solely to the elevated risk of fertilizer runoff caused by high slope values, but concurrently to their poor moisture content.

VI. Conclusion

The determination of fertilizer management zones allows for the design of targeted and optimized treatment strategies. While existing methods consider various factors, none explicitly address the crucial aspect of fertilizer responsivity in

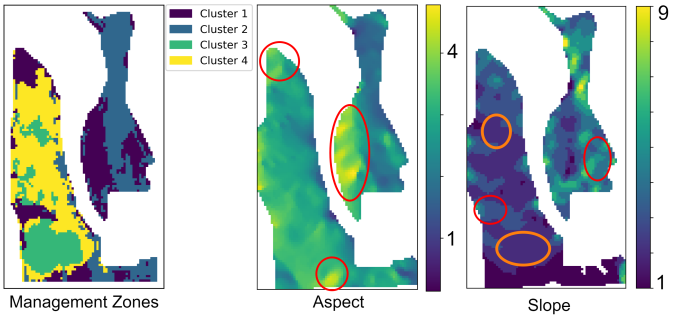


Fig. 10: Zones obtained for field A along its A and S maps.

defining these zones. To address this gap, we introduce a novel management zone clustering method based on neural network-generated response curves that accounts for the within-field variability of fertilizer responsivity.

Our approach leverages N-response curves generated using a specialized 2D regression convolutional neural network, providing an approximation of how crop yield responds to varying fertilizer rates. We characterize the shape dissimilarity of these curves and use it as a metric for clustering analysis. Experimental results on two winter wheat dryland fields demonstrate the effectiveness of our method in organizing field points into MZs with consistent N-response curve shapes. Thus, each MZ exhibits a distinct fertilizer responsivity pattern.

Finally, our counterfactual explanation method improves understanding of the impact of various covariate factors on the assignment of field points to specific MZs. Our experiments provide explanations suited to each MZ and field, uncovering the influence of specific terrain characteristics on fertilizer responsivity. Future work will focus on the design of statistical tests that would allow for determining the optimal number of MZs. In the context of On-Farm Precision Experimentation (OFPE), we also plan to modify our approach to assist in creating optimal fertilizer maps for the different zones. Specifically, instead of using the conventional rectangular gridding, we will employ several small, homogeneous regions that align better with the local variations in the field.

Acknowledgements

The authors wish to thank the team members of the On-Field Precision Experiment (OFPE) project for their comments throughout the development of this work. This research was supported by a USDA-NIFA-AFRI Food Security Program Coordinated Agricultural Project (Accession Number 2016-68004-24769), and also by the USDA-NRCS Conservation Innovation Grant from the On-farm Trials Program (Award Number NR213A7500013G021).

References

[1] R. Khosla, K. Fleming, J. A. Delgado, T. M. Shaver, and D. G. Westfall, "Use of site-specific management zones to improve nitrogen management for precision agriculture," *Journal of Soil and Water Conservation*, vol. 57, no. 6, pp. 513–518, 2002.

[2] R. B. Ferguson, G. W. Hergert, J. S. Schepers, C. A. Gotway, J. E. Cahoon, and T. A. Peterson, "Site-specific nitrogen management of irrigated maize," *Soil Science Society of America Journal*, vol. 66, no. 2, pp. 544–553, 2002.

[3] N. Davatgar, M. Neishabouri, and A. Sepaskhah, "Delineation of site-specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering," *Geoderma*, vol. 173-174, pp. 111–118, 2012.

[4] D. B. Jaynes, T. C. Kaspar, T. S. Colvin, and D. E. James, "Cluster analysis of spatiotemporal corn yield patterns in an Iowa field," *Agronomy Journal*, vol. 95, no. 3, pp. 574–586, 2003.

[5] A. Ali, R. Martelli, E. Scudiero, L. F., F. G., R. V., and L. Barbanti, "Soil and climate factors drive spatio-temporal variability of arable crop yields under uniform management in northern Italy," *Archives of Agronomy and Soil Science*, vol. 69, no. 1, pp. 75–89, 2023.

[6] C. Georgi, D. Spengler, S. Itzerott, and B. Kleinschmit, "Automatic delineation algorithm for site-specific management zones based on satellite remote sensing data," *Precision Agriculture*, vol. 19, no. 4, p. 684–707, Nov. 2017.

[7] D. J. Gallardo-Romero, O. E. Apolo-Apolo, J. Martínez-Guanter, and M. Pérez-Ruiz, "Multilayer data and artificial intelligence for the delineation of homogeneous management zones in maize cultivation," *Remote Sensing*, vol. 15, no. 12, 2023.

[8] J. Reyes, O. Wendroth, C. Matocha, and J. Zhu, "Delineating site-specific management zones and evaluating soil water temporal dynamics in a farmer's field in Kentucky," *Vadose Zone Journal*, vol. 18, no. 1, p. 180143, 2019.

[9] B. Shashikumar, S. Kumar, K. George, and A. Singh, "Soil variability mapping and delineation of site-specific management zones using fuzzy clustering analysis in a Mid-Himalayan watershed, india," *Environment, Development and Sustainability*, vol. 25, no. 8, p. 8539–8559, 2022.

[10] A. Ali, V. Rondelli, R. Martelli, G. Falsone, F. Lupia, and L. Barbanti, "Management zones delineation through clustering techniques based on soils traits, NDVI data, and multiple year crop yields," *Agriculture*, vol. 12, no. 2, 2022.

[11] S. Maleki, A. Karimi, A. Mousavi, R. Kerry, and R. Taghizadeh-Mehrjardi, "Delineation of soil management zone maps at the regional scale using machine learning," *Agronomy*, vol. 13, no. 2, 2023.

[12] S. H. Javadi, A. Guerrero, and A. M. Mouazen, "Clustering and smoothing pipeline for management zone delineation using proximal and remote sensing," *Sensors*, vol. 22, no. 2, p. 645, Jan. 2022.

[13] J. J. Fridgen, N. R. Kitchen, K. A. Sudduth, S. T. Drummond, W. J. Wiebold, and C. W. Fraisse, "Management zone analyst (mza)," *Agronomy Journal*, vol. 96, no. 1, pp. 100–108, 2004.

[14] G. Morales, J. Sheppard, A. Peerlinck, P. Hegedus, and B. Maxwell, "Generation of site-specific nitrogen response curves for winter wheat using deep learning," in *15th International Conference on Precision Agriculture*, 2022.

[15] G. Morales and J. Sheppard, "Counterfactual explanations of neural network-generated response curves," in *International Joint Conference on Neural Networks*, 2023, pp. 01–08.

[16] G. Morales, J. W. Sheppard, P. B. Hegedus, and B. D. Maxwell, "Improved yield prediction of winter wheat using a novel two-dimensional deep regression neural network trained via remote sensing," *Sensors*, vol. 23, no. 1, p. 489, jan 2023.

[17] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[18] M. Zeraatpisheh, E. Bakhshandeh, M. Emadi, T. Li, and M. Xu, "Integration of PCA and fuzzy clustering for delineation of soil management zones and cost-efficiency analysis in a citrus plantation," *Sustainability*, vol. 12, no. 14, 2020.

[19] S. Barocas, A. D. Selbst, and M. Raghavan, "The hidden assumptions behind counterfactual explanations and principal reasons," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 2020, p. 80–89.

[20] R. Kommiya, D. Mahajan, C. Tan, and A. Sharma, "Towards unifying feature attribution and counterfactual explanations: Different means to the same end," in *Conf. on AI, Ethics, and Society*, 2021, p. 652–663.

[21] J. Ramsay and B. Silverman, *Functional Data Analysis*. Springer, 2005.