The New England Journal of Statistics in Data Science Volume 3, 82–93 (2025)
DOI: https://doi.org/10.51387/24-NEJSDS71

Irrigation Zone Delineation by Coupling Neural Networks with Spatial Statistics

Matthew J. HEATON*, David TEUSCHER, and Neil C. HANSEN

Abstract

Variable rate irrigation (VRI) seeks to increase the efficiency of irrigation by spatially adjusting water output within an agricultural field. Central to the success of VRI technology is establishing homogeneous irrigation zones. In this research, we propose a fusion of statistical modeling and deep learning by using artificial neural networks to map irrigation zones from simple-to-measure predictors. We further couple our neural network model with spatial correlation to capture smooth variations in the irrigation zones. We demonstrate the effectiveness of our model to define irrigation zones for a farm of winter wheat crop in Rexburg, Idaho.

KEYWORDS AND PHRASES: Ordered multinomial, Precision agriculture, Bayesian.

1. INTRODUCTION

1.1 Problem Background and Data

The management of agricultural fields (i.e. farming) in the era of data science has evolved to use spatial mapping, remote sensing, soil and terrain measurements, weather measurements and other data sources to improve the quantity and quality of crops [21, 9]. The management of agricultural fields using advanced data analytics is referred to, collectively, as "precision agriculture" [5, 8] and is quickly becoming industry standard. Broadly, precision agriculture attempts to use the spatial variability within a field to manage individual crop locations rather than treat a field as spatially and temporally homogeneous. The spatial variability within a field can be influenced by differences in elevation, aspect, or other topographical features of the field that influence both irrigation and crop yield. Given that agriculture fields often occupy multiple acres of space, precision agriculture has been shown to outperform basic farming techniques by increasing crop yield [30, 32].

Crop yield is heavily driven by the volumetric water content (VWC; the ratio of the volume of water currently in the soil to the unit volume of soil, or in other words the quantity of water contained in the soil) and, in arid regions, irrigation is a practiced method for controlling and adjusting the VWC. As such, variable rate irrigation (VRI) is a practice within precision agriculture focused on using data to adjust the amount of water applied throughout the field according to spatial and temporal variations [20, 39, 40, 25, 60]. Such management zones can be temporally dynamic to help inform day-to-day watering decisions [13] or, as is the case in this application, static to inform season-long watering decisions.

One approach to VRI is to partition the field into management zones (or irrigation zones) wherein irrigation rates are homogeneous within each management zone but different across zones rather than utilizing a constant irrigation rate throughout the entire field [18, 61, 57, 38, 37]. Benefits of using VRI with management zones include reductions in water and energy input, increased crop productivity, decreased runoff (i.e. less waste) and reductions in chemical inputs and soil pollution [42, 3, 25].

Effective irrigation zones should partition the agricultural field based on VWC [31]. However, obtaining VWC is a labor intensive process and is often only sparsely measured for the entire field [27]. As an example, consider Figure 1 which shows 66 VWC measurements averaged over four time periods between April and September 2019 for an agricultural field of winter wheat in Rexburg, Idaho along with the associated irrigation zone for each location. Although the VWC was recorded at 66 different locations, the available data do not adequately cover the field to the point where precise irrigation decisions and zones can be precisely defined. That is, based on such sparse measurement the boundaries between zones are unclear.

Rather than using the VWC directly, agricultural scientists have begun to use alternative, more easily obtained, data to inform irrigation zones [50, 56, 26, 24]. Examples of such possible data for the Rexburg field considered in this research are shown in Figure 2 and include historical yields, normalized difference vegetation index (NDVI), elevation, topographical wetness index (TWI), aspect, and slope (more details on these covariates are given in the application Section 3.1 below). These alternative covariates to VWC were recorded for over 5000 unique locations in the field using

*Corresponding author.

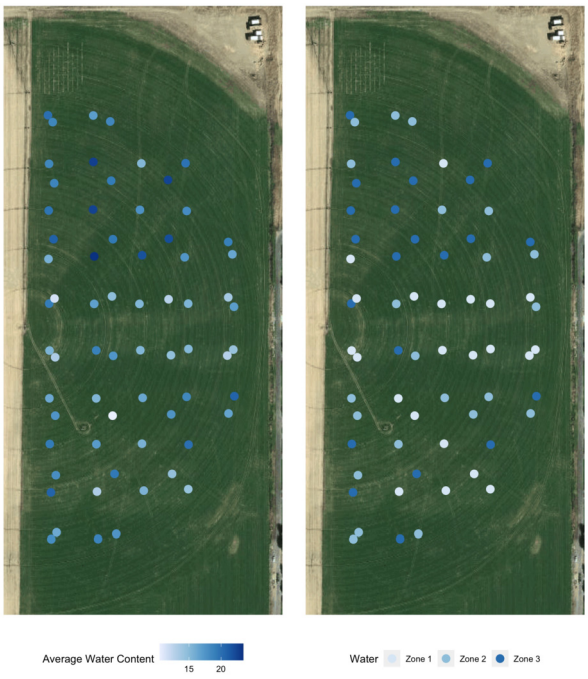


Figure 1: The scatterplot on the left shows the average VWC at the 66 samples throughout the field in Rexburg, Idaho. The edges of the field were excluded from this study. The scatterplot on the right shows the classified zone at the 66 samples throughout the field in Rexburg, Idaho. The average VWC was calculated from 4 VWC measurements taken between April and August of the same year.

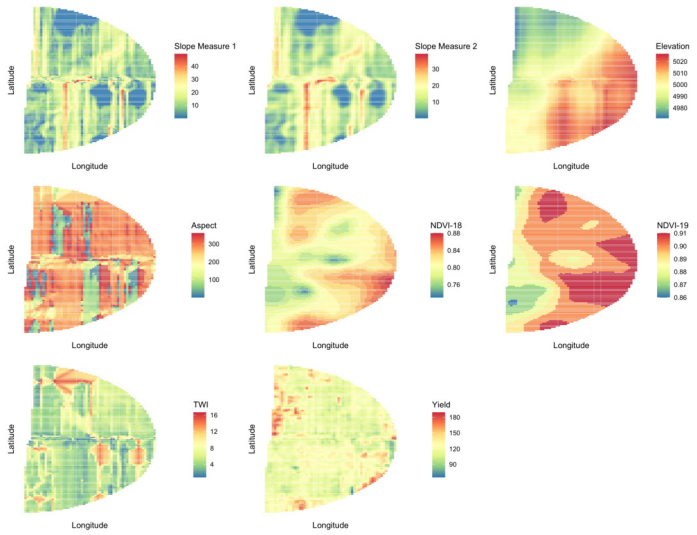


Figure 2: Heat maps for the alternative data to VWC across the field in Rexburg using remote sensing.

simple remote sensing mounted on drones or tractors making these data readily available to farmers.

The primary issue of using alternative covariates to VWC such as that in Figure 2 to define VRI zones is that such covariates are non-linearly related to VWC and, hence, may result in less water efficient zones than ones defined directly by VWC. Figure 3 shows scatterplots of a few of the variables against the VWC for the Rexburg field demonstrating clear non-linear relationships. In fact, the scatterplots in Figure 3 show a highly complex relationship between these alternative covariates and VWC suggesting that zones defined directly on the covariates may vary considerably from the more water-efficient zones defined on VWC.

1.2 Research Goals and Contributions

In this research, we seek to implement a method to delineate irrigation zones based on VWC using the data displayed in Figure 2 as covariate information. Specifically, we define a model relating the easily obtained and spatially dense covariate data to the sparse VWC data. To do so, we integrate deep learning into a statistical modeling framework for irrigation zones to capture the complex relationship between the covariate data and the response variable.

Deep learning is a subfield of machine learning focused on using highly flexible algorithms called neural networks to learn complex relationships between covariates and response variables [53]. Deep learning algorithms have been successfully implemented across a wide range of applications leading to their explosion in popularity in recent years [1, 41, 49, 12]. More recently, deep learning (and machine learning more generally) have successfully incorporated spatial correlation [48, 45] into these algorithms.

From a statistical perspective, deep learning algorithms do not inherently account for uncertainty in the predictions and are often so complex that they are uninterpretable (so-called "black box" methods). Yet, recent efforts in the statistics community have started to fold deep learning into statistical models [34, 54, 46, 11]. While we acknowledge that some methods (such as dropout) can give uncertainty estimates [see 15], we consider embedding a deep neural network into a statistical model and follow [52] by estimating associated parameters using a fully Bayesian paradigm. Further, we implement Bayesian versions of partial dependence plots to give some interpretability to the deep neural network between the covariates and the response variable.

With a deep neural network accounting for the complex relationships between the covariates and response, our model also creates smooth irrigation zones by incorporating spatial correlation. Specifically, we use spatial basis functions distinctly from the neural network portion of the model thereby successfully merging deep learning into a spatial modeling framework.

The primary quantity of interest in this research is the irrigation zone of each location on the field – not necessarily the VWC. Admittedly, we could predict the VWC first and then create irrigation zones by splitting based on predicted VWC quantiles. However, here we perform modeling for the associated zone as the response of interest given by the right panel of Figure 1. We focus on the zone rather than VMC

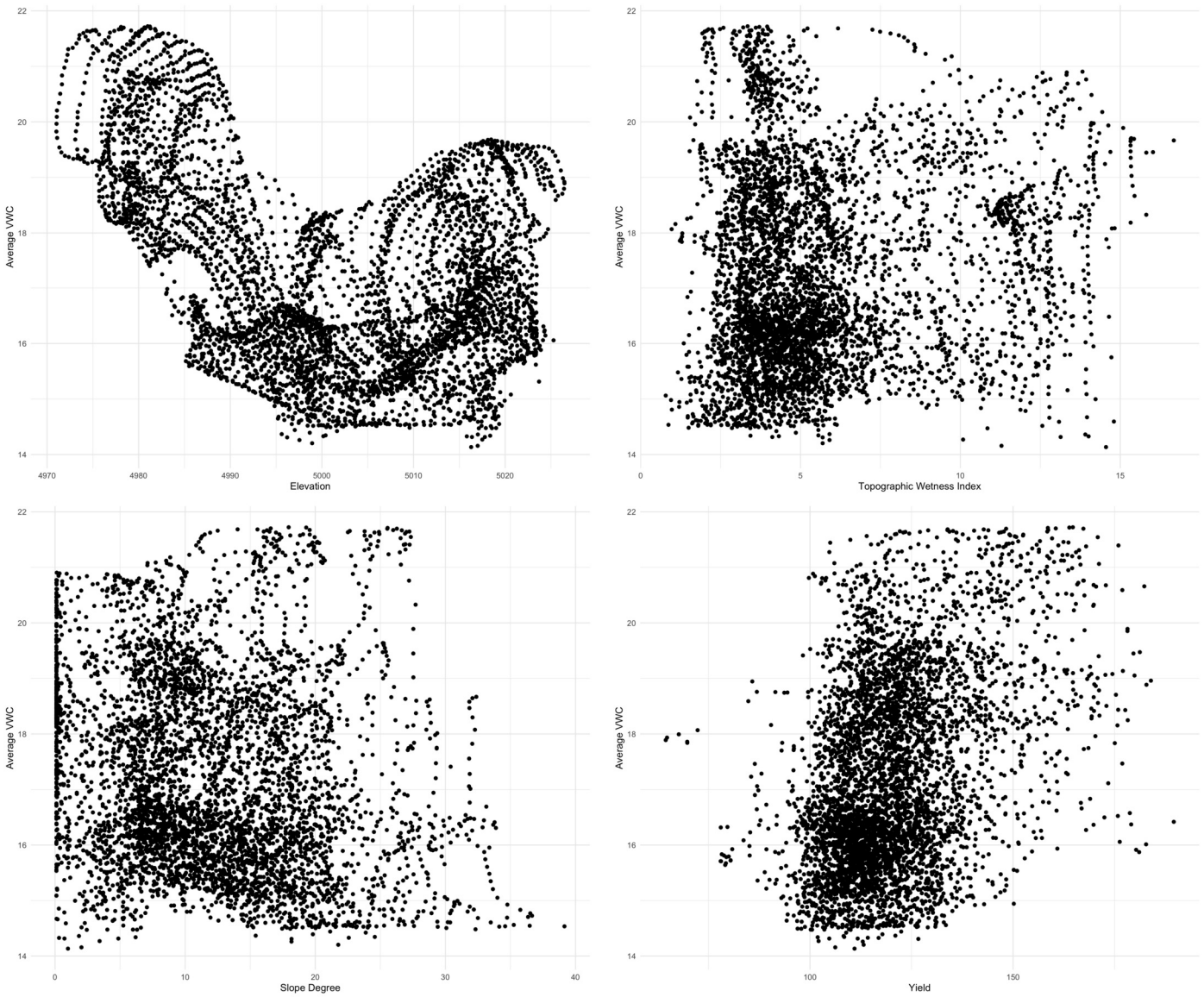


Figure 3: Scatterplots of the water content against elevation, TWI, slope, and yield throughout the field. It is evident, especially for elevation and slope, that the relationship between these covariates and the water content is non-linear, justifying the use of neural networks to model this relationship.

for a few reasons. First, our data is unique in that we have the exact VWC measurements. However, given the cost, obtaining exact VWC data is exceptionally rare in agricultural practice (our data being an exception rather than the rule). Predicting continuous VWC would be applicable only to this field in Rexburg and would not be able to be tested on other fields. Rather, most agriculture fields are able to only collect ambiguous VWC levels of "low" or "high." By focusing on irrigation zone, our statistical model will be more aligned with typical data available in other crop fields. Second, VRI technology is typically done by applying "less than average" water to some locations and "more than average" to others. Hence, what is required to implement VRI is knowledge of which locations are "low" and which are "high" rather than the exact VWC.

Given the focus on irrigation zones, the response variable is thus an ordered multinomial response. Under this discretization, we build our model following the latent variable approach of [2], [22] and [4] to facilitate Bayesian computation. While others have developed statistical models for neural networks [see 16, 34], to our knowledge, this is the first attempt at building a deep spatial statistical model for an ordered categorical response.

We view this work as a contribution to the growing body of literature of spatial deep learning methods [see 58]. Alternative approaches include [7] use spatial bases as inputs into

neural networks. However, here we use spatial bases separately from the neural network because using such bases as inputs may sacrifice some spatial smoothness which is crucial to our application. Likewise, [59] use generalized squared error loss functions with a spatial covariance to fit spatial graphical neural networks. Our approach is inherently different from this because our response is non-Gaussian so squared error loss functions are not directly applicable.

The remainder of this paper is outlined as follows. Section 2 describes our spatial neural network model along with our Bayesian strategy to estimate model parameters. Section 3 describes the result of the model fit to the Rexburg field including model tuning, predicted zone delineation and effect interpretation. Finally, Section 4 concludes by pinpointing strengths and weaknesses of our approach along with areas for future research.

2. A SPATIAL NEURAL NETWORK MODEL

2.1 Model Specification

Let $Y(\boldsymbol{s}) \in \{1, \ldots, R\}$ denote the irrigation zone for location $\boldsymbol{s} = (s_1, s_2)' \in \mathbb{R}^2$ where R is the number of desired irrigation zones for a field. Given the current state of variable rate irrigation systems, usually R will only be as high as 4 or 5. Further, because each irrigation zone is given a different amount of water (i.e. dry irrigation zones receive more water), we can treat $Y(\boldsymbol{s})$ as an *ordered*, discrete spatial random variable. Due to the difficulty of performing spatial analysis on a discrete scale, we follow [2], [22] and [4] by augmenting $Y(\boldsymbol{s})$ with a latent variable $Z(\boldsymbol{s}) \in \mathbb{R}$ such that

$$\big(Y(\boldsymbol{s}) \mid \{c_r\}, Z(\boldsymbol{s})\big) = \sum_{r=1}^{R} r \mathbb{1}\big(c_{r-1} \leq Z(\boldsymbol{s}) < c_r\big) \qquad (2.1)$$

where $\mathbb{1}(\cdot)$ is an indicator function and $c_0 = -\infty < c_1 = 0 < c_2 < \cdots < c_R = \infty$ are cut points used to determine the probability that location $\boldsymbol{s}$ belongs to a specific zone.

Under Equation (2.1) and given the cut points $c_0, \ldots, c_R$, $Y(\boldsymbol{s})$ is completely determined by $Z(\boldsymbol{s})$ so that modeling $Z(\boldsymbol{s})$ will induce a statistical model for $Y(\boldsymbol{s})$. Using this relationship between $Y(\boldsymbol{s})$ and $Z(\boldsymbol{s})$, we assume

$$\big(Z(\boldsymbol{s}) \mid w(\boldsymbol{s})\big) \sim \mathcal{N}\big(f_L\big(\boldsymbol{X}(\boldsymbol{s})\big) + w(\boldsymbol{s}), 1\big) \qquad (2.2)$$

where $f_L(\boldsymbol{X}(\boldsymbol{s}))$ is the univariate output function of an L-layer feed forward neural network (FFNN) or, alternatively referred to as a multilayer perceptron (MLP) with input covariates $\boldsymbol{X}(\boldsymbol{s}) = (X_1(\boldsymbol{s}), \ldots, X_P(\boldsymbol{s}))'$ and $w(\boldsymbol{s})$ is a spatial random effect. Under this model structure, we follow [2] by fixing the variance at 1 to identify the scale of the weights of the neural networks (coefficients) and we set $c_1 = 0$ to identify a bias term. This model, importantly, accounts for the various challenges of determining irrigation zones via statistical modeling discussed in Section 1.2. First, the FFNN in (2.2) utilizes a nonlinear relationship between the set of covariates $\boldsymbol{X}(\boldsymbol{s})$ and the associated irrigation zone (see Figure 3). Second, the spatial random effect $w(\boldsymbol{s})$ in (2.2) ensures that the fitted irrigation zones change smoothly over space thereby creating zones that can be reasonably implemented via VRI. Consider each piece of the model in the following paragraphs.

FFNNs have become known for the strong performance in prediction, especially where non-linear relationships exist between covariates and the response variable as is the case here [12]. A FFNN with L layers, consists of three different types of layers: an input layer, hidden layers and an output layer each consisting of a different number of units. More succinctly, let $\boldsymbol{f}_l(\boldsymbol{X}(\boldsymbol{s})) = (f_{\ell 1}(\boldsymbol{X}(\boldsymbol{s})), \ldots, f_{\ell P_\ell}(\boldsymbol{X}(\boldsymbol{s})))'$ be the vector of P_ℓ units at layer ℓ of the FFNN. The transformation from layer to layer occurs via

$$\boldsymbol{f}_\ell\big(\boldsymbol{X}(\boldsymbol{s})\big) = a_\ell\big(\boldsymbol{\lambda}_{(\ell-1)0} + \boldsymbol{\Lambda}_{\ell-1}\boldsymbol{f}_{\ell-1}\big(\boldsymbol{X}(\boldsymbol{s})\big)\big) \qquad (2.3)$$

where $a_\ell(\cdot)$ is a element-wise nonlinear activation function used at layer ℓ, $\boldsymbol{\lambda}_{(\ell-1)0} = (\lambda_{(\ell-1)01}, \ldots, \lambda_{(\ell-1)0P_\ell})'$ is the vector of intercepts (biases) applied to layer $\ell - 1$ and $\boldsymbol{\Lambda}_{\ell-1} = \{\lambda_{(\ell-1)ij}\}_{i,j}$ is the $P_\ell \times P_{\ell-1}$ matrix of coefficients (weights) used to transition from layer $\ell - 1$ to layer ℓ. For the input layer (i.e. $\ell = 1$), we take $\boldsymbol{f}_1(\boldsymbol{X}(\boldsymbol{s})) = \boldsymbol{X}(\boldsymbol{s})$ such that $P_1 = P$ is the number of covariates included in the model. From a statistical point of view, the weights and biases (intercepts and coefficients) $\{\boldsymbol{\lambda}_{\ell 0}, \boldsymbol{\Lambda}_\ell\}_{\ell=1}^{L-1}$ are the model parameters to be estimated from the data while the number of layers L, the layer dimensions $\{P_\ell\}$ and the activation functions are known as "tuning parameters" and are fixed via cross-validation to optimize prediction performance.

The flexibility of FFNNs to capture complex relationships comes from appropriate choice of L, $\{P_\ell\}$ and $\{a_\ell(\cdot)\}$. First, deeper networks (larger L) may be required to allow for more complex transformations of $\boldsymbol{X}(\boldsymbol{s})$ between the input and output layers. Conversely, in the simplest case, when $L = 2$, the relationship between $\boldsymbol{X}(\boldsymbol{s})$ and $Z(\boldsymbol{s})$ would be linear provided $a_L(x) = x$ is the identity function.

Beyond the number of layers, the number of units per layer (denoted by P_ℓ) can also be modified to capture complex relationships between the covariates and the response variable. For example, if $P_\ell > P_1$ for any $\ell > 1$ then the FFNN uses a dimension expansion to model the relationship between covariates and the response [55, 6, 47]. As recently shown by [28] and [33] but pioneered by [35], an infinite dimension expansion would mimic a Gaussian process regression model between the covariates and response. Conversely, if $P_\ell < P_1$ for any $\ell > 1$ then the transition in (2.3) acts similar to a principal components regression model by reducing the dimension of the input space [see 14, Chapter 5].

The activation functions $\{a_\ell(\cdot)\}_{l=2}^{L}$ in (2.3) are typically non-linear to ensure that the relationship between $\boldsymbol{X}(\boldsymbol{s})$ and $Z(\boldsymbol{s})$ is also non-linear. Typical choices of $a_\ell(\cdot)$ include rectified linear units, hyperbolic tangents and identity (see [36]

and [44] for a discussion of these activations). However, simple algebraic manipulations show that if $a_\ell(\boldsymbol{x}) = \boldsymbol{x}$ (the identity activation) for all ℓ then the relationship between $Z(\boldsymbol{s})$ and $\boldsymbol{X}(\boldsymbol{s})$ would be linear. Notably, to increase flexibility, different layers can utilize different activation functions but, traditionally, the same activation function is used for all layers with the exception of the output layer where the domain of $a_L(\cdot)$ must match the support of $Z(\boldsymbol{s})$ (which, in our cases, $a_L(\cdot)$ is the identity function).

While $f_L(\boldsymbol{X}(\boldsymbol{s}))$ captures the non-linear relationship between the response and covariates, the spatial random effect $w(\boldsymbol{s})$ serves to smooth the predicted zones over space. To achieve this smoothing, we opt to use the Moran basis function expansion for $w(\boldsymbol{s})$ advocated by, among others, [23]. To calculate the Moran basis functions, we first determine an adjacency matrix $\boldsymbol{A} = \{a_{ij}\}$ between all observed locations and desired prediction locations according to the inverse weighting distance

$$a_{ij} = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|} & \text{otherwise} \end{cases} \tag{2.4}$$

where $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ are two locations on the field (either observed or predicted). Let $\boldsymbol{P} = \boldsymbol{I} - \mathbf{1}\mathbf{1}'/n$ where $\mathbf{1}$ is a vector of ones and n is the total data size (number of observed locations plus the number of desired prediction locations). Let $\boldsymbol{B}$ be the $n \times K$ matrix of the K eigenvectors associated with the K largest positive eigenvalues of $\boldsymbol{PAP}$. Then, we set

$$w(\boldsymbol{s}) = \boldsymbol{b}'(\boldsymbol{s})\boldsymbol{\beta} \tag{2.5}$$

where $\boldsymbol{b}'(\boldsymbol{s})$ the row of $\boldsymbol{B}$ associated with location $\boldsymbol{s}$ and $\boldsymbol{\beta}$ is a vector of coefficients to be estimated from the data. Generally, as K increases then the associated fitted spatial surface from the Moran basis is more variable. Hence, for purposes of this research, we treat K as an additional tuning parameter that we will choose via cross-validation.

2.2 Parameter Estimation

Even though neural networks are typically fit via loss minimization, in an effort to merge machine learning and statistical modeling, we adopt a Bayesian approach for parameter estimation to ensure that our predicted irrigation zone surface accounts for associated parameter uncertainty. Accounting for uncertainty here is important so that the predicted irrigation zones can be potentially altered to more easily be incorporated into a VRI system. For example, if the zone assigned to a certain location is highly uncertain, then that location can be manually assigned to a zone by the farmers to increase the overall efficiency of the VRI.

Under the Bayesian approach, prior distributions are required for all the neural network parameters $\{\boldsymbol{\lambda}_{\ell 0}, \boldsymbol{\lambda}_\ell\}$, the Moran basis coefficients ($\boldsymbol{\beta}$) and the cut points $c_2, \ldots, c_R$ and parameter estimation is done via posterior inference. Because the $\boldsymbol{\beta}$ parameter vector is simply coefficients in an ordered probit regression model, we assume a vague $N(0, 100\boldsymbol{I})$ prior distribution because the data can generally estimate these parameters well [22]. Because the cut points are ordered so that $0 < c_2 < \cdots < c_R = \infty$, each cutpoint was transformed according to $c_2^\star = \log(c_2)$ and

$$c_r^\star = \log(c_r - c_{r-1}) \tag{2.6}$$

for $r = 3, \ldots, R-1$ and a $\mathcal{N}(0, 10)$ prior was used for each $c_r^\star$. The corresponding back transformation $c_r = \sum_{i=2}^{r} \exp\{c_i^\star\}$ ensures the ordering constraint.

The priors for the FFNN weights $\{\Lambda_\ell\}$ and biases $\{\lambda_{0\ell}\}$ need to be chosen with care to avoid overfitting. As documented by, among others, [10], neural networks can easily overfit training data resulting in poor predictive performance. One common approach for restricting neural networks is via penalization (regularization). In a Bayesian setting, regularization is enforced via informative prior constraints [see 43, for a review]. As such, we assume *a priori* independent $\mathcal{N}(0, 0.01)$ priors for all biases and weights. The highly informative 0.01 variance constrains the biases and weights to be near zero; thus preventing overfitting similar to ridge and LASSO regression models. For our application below, we tried a few different values for the prior variance here but ultimately settled on 0.01 as the best performing prior in terms of predictive accuracy (see Section 3.1 below). Admittedly, a Laplace prior could be used for the weights because it corresponds to the LASSO penalty [51]. For our purposes, the above Gaussian prior worked adequately to prevent overfitting as demonstrated by our cross-validation study below.

Posterior inference for our model parameters was accomplished by using Markov chain Monte Carlo (MCMC) sampling. Conditional on all other parameters, the complete conditional distribution for $Z(\boldsymbol{s})$ is a truncated Gaussian distribution with mean $f_L(\boldsymbol{X}(\boldsymbol{s})) + \boldsymbol{b}'(\boldsymbol{s})\boldsymbol{\beta}$, variance 1 and endpoints $c_{Y(\boldsymbol{s})-1}$ and $c_{Y(\boldsymbol{s})}$. Because each $Z(\boldsymbol{s})$ is conditionally independent, this sampling can be done independently and efficiently. Next, because $Z(\boldsymbol{s}) \in \mathbb{R}$, we use assume $a_L(\cdot)$ is the identity activation and $P_L = 1$ so that

$$Z(\boldsymbol{s}) \overset{iid}{\sim} \mathcal{N}\big(\lambda_{0(L-1)} + \boldsymbol{f}'_{L-1}\big(\boldsymbol{X}(\boldsymbol{s})\big)\boldsymbol{\Lambda}_{L-1} + \boldsymbol{b}'(\boldsymbol{s})\boldsymbol{\beta}, 1\big) \tag{2.7}$$

which can be rewritten simply as $Z(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{X}'_\star(\boldsymbol{s})\boldsymbol{\beta}^\star, 1)$ where $\boldsymbol{\beta}^\star = (\lambda_{0(L-1)}, \boldsymbol{\Lambda}'_{L-1}, \boldsymbol{\beta}')'$ and $\boldsymbol{X}_\star(\boldsymbol{s}) = (1, \boldsymbol{f}'_{L-1}(\boldsymbol{X}(\boldsymbol{s})), \boldsymbol{b}'(\boldsymbol{s}))'$. Under the above Gaussian priors, the complete conditional for $\boldsymbol{\beta}^\star$ is Gaussian and can be sampled directly.

While $\{Z(\boldsymbol{s})\}$ and $\{\lambda_{0(L-1)}, \boldsymbol{\Lambda}_{L-1}, \boldsymbol{\beta}\}$ can be sampled directly from the corresponding complete conditional distributions, the cut points $c_2, \ldots, c_{R-1}$ and other FFNN parameters $\{\boldsymbol{\lambda}_{0l}, \boldsymbol{\Lambda}_l\}_{l=1}^{L-2}$ need to be sampled indirectly via Metropolis-Hastings or other algorithms. In early phases of this research, the weights and biases were sampled individually or a single column at a time. However, this proved to

be computationally expensive and there was high correlation between the different weights within each layer. As a result, our final MCMC algorithms sampled the neural net weight matrix $\mathbf{\Lambda}_\ell$ and the biases $\boldsymbol{\lambda}_{0\ell}$ were sampled jointly for each layer (with each layer being sampled separately). For our MCMC algorithm, we used the adaptive Metropolis algorithm from [17] to update the proposal variance and achieve better mixing for the FFNN parameters. Finally, this adaptive Metropolis algorithm was again used to sample the transformed cutpoints $c_2^\star, \ldots, c_{R-1}^\star$ after marginalizing out $Z(\boldsymbol{s})$ as suggested by [22].

3. APPLICATION TO REXBURG FIELD

3.1 Model Settings and MCMC Diagnostics

For our application, we used ten covariates: elevation, yield, NDVI index for 2018 and 2019, two different measures of slope at a location, the x and y aspect of a location and a topographical wetness index (TWI). Each of the covariates were observed at 5062 10 m^2 areas in the field (see Figure 2). In this application, NDVI (a measure of vegetation greenness) was calculated using an aerial drone at four distinct times throughout a growing season and averaged as a single NDVI measurement for the year. Yield was measured per 10 m^2 unit of land using a grain yield monitor installed on the harvester. TWI, is a static measure, calculated from a digital elevation models, and indicates where water will accumulate in an area with elevation variability. Slope as well as the x and y aspects were likewise calculated from a digital elevation model.

The data of irrigation zone consisted of only the 66 observations shown in the right panel of Figure 1. For our implementation, we set the number of irrigation zones, R, to 3. As previously mentioned, R will rarely if ever be higher than 4 or 5 and, in discussion with the farm owner, 3 zones was determined to be reasonable based on the field's VRI capacities. We also used rectified linear unit activation functions for all the activation functions with the exception of the output layer which was an identity activation to match the support of $Z(\boldsymbol{s})$. Certainly, other activation functions could be used but the rectified linear units is one of the most common.

Beyond the above model settings, implementation of our spatial neural network model requires tuning the number of layers (L), the number of units per layer ($\{P_\ell\}$), the number of Moran basis functions (K) and the penalization for the weights in the prior (the prior variance). Computationally, fitting the neural network model at a single setting took about 3.5 hours on a 2.60 GHz CPU. For each parameter setting in Table 1 and for each of 0.1, 0.01 and 0.001 for the prior variance of the weights, we implemented a 6-fold cross validation and calculated the average adjusted Rand index (ARI). The ARI is given by

Table 1. Cross Validation results. The first column lists the number of neurons for each layer, the second column indicates the number of layers for the neural net, the third column shows the number of Moran basis functions, and the fourth column gives the average adjusted Rand index over the 6 folds.

Neurons	Layers	Spatial	Rand index
10	1	0	.06767
(10,10)	2	0	.1776
20	1	0	.12296
10	1	5	.31795
(10,10)	2	5	.22377
20	1	5	.13879

$$ARI = \frac{\sum_{ij}\binom{n_{ij}}{2} - [\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}]\Big/\binom{n}{2}}{\frac{1}{2}[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}] - [\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}]\Big/\binom{n}{2}}$$

where a_r is the number locations predicted to belong to irrigation zone r and b_r is the number of true locations belonging to irrigation zone r for $r = 1, \ldots, R$ and n_{ij} is number of location classified as zone i in the predictions and zone j in the truth. ARI, intuitively, is a measure of similarity between two data clusterings (in our case, the similarity between the true zone and the predicted zone) and ranges from -1 to 1 where 1 indicates perfect agreement between the two clusterings, 0 indicates a random agreement and -1 indicates that the two clusterings are completely different. Prior to fitting the models in Table 1, a larger grid was first used to get a general idea of the reasonable values for the tuning parameters. Given our data set consisted of only 66 observations, an effort was made to keep the total number of the parameters low. Deeper neural networks, generally, require big data to be effective. Hence, the grid search only examines one or two layer neural networks (in addition to the input and output layers). The maximum number of Moran basis functions considered was 10 in order to make sure that the zones were being determined by the neural network predictions rather than being overly driven by the spatial aspect of our model.

Table 1 shows the cross validation results using the 0.01 prior variance because this value was uniformly better in terms of predictive accuracy. Cross-validation finds that a single hidden layer with 10 neurons and 5 Moran basis functions were the ideal parameters for this data. The results displayed in the following subsections are from this model setting. Note that, generally, from Table 1, adding spatial basis functions improved the model's predictive ability suggesting that merging spatial modeling techniques with deep learning is of value in this particular setting.

When using Markov chain Monte Carlo sampling in a Bayesian framework as is the case here, it is important to

make sure that the algorithm provides samples of the parameters from posterior distribution via convergence diagnostics. The supplementary material to this article includes trace plots of $f_L(\boldsymbol{X}(\boldsymbol{s})) + w(\boldsymbol{s})$ for four different locations along with a trace plot of an example cut point. We focus on these trace plots as $f_L(\boldsymbol{X}(\boldsymbol{s})) + w(\boldsymbol{s})$ and the cut point because these are the main quantities from which we derive our irrigation zone delineation. These trace plots show that these parameters have converged and they can be used to for posterior inference.

3.2 Model Comparisons

As a means to validate the use of the spatial neural network model proposed in this analysis, three additional alternative models were fit to the data and the ARI was again calculated for the 66 locations across the field in Rexburg. The four models compared are ordinal logistic regression with linear effects (Logistic), ordinal logistic regression with natural splines (NSLogistic), ordinal logistic regression with natural splines and spatial bases (Spatial NSLogistic), the neural network model with no spatial basis functions (NN) and our full neural network model with spatial basis functions (Spatial NN). Ordinal logistic regression is a common approach to modeling discrete ordinal multinomial response variables. One of the drawbacks is the assumptions that the relationship between predictors and the response is linear. Splines are a common approach to modeling non-linear relationships for statisticians, while neural networks are often a machine learning approach to account for these non-linear relationships. These two models can provide a comparison between these two different approaches to model a non-linear relationship. While neural networks and splines are able to account for the non-linear relationships, they don't incorporate any spatial correlation. For irrigation zone delineation, the spatial correlation is used to smooth out the irrigation zone predictions across the field. As a result, we also included the spatial basis functions beyond the natural splines and/or neural network to help improve the predictions of the zones.

Table 2 shows the adjusted Rand index for the four models using all 66 locations in the field. The results show that the ordinal logistic regression model is least able to delineate zones. This is, perhaps, not surprising given this model doesn't account for the non-linear relationships between predictors and the response variable or the spatial correlation.

Table 2. RAND index for the four different models for the 66 locations in the field.

Model	ARI
Logistic	0.2065
NSLogistic	0.2273
Spatial NSLogistic	0.2412
NN	0.2245
Spatial NN	0.2741

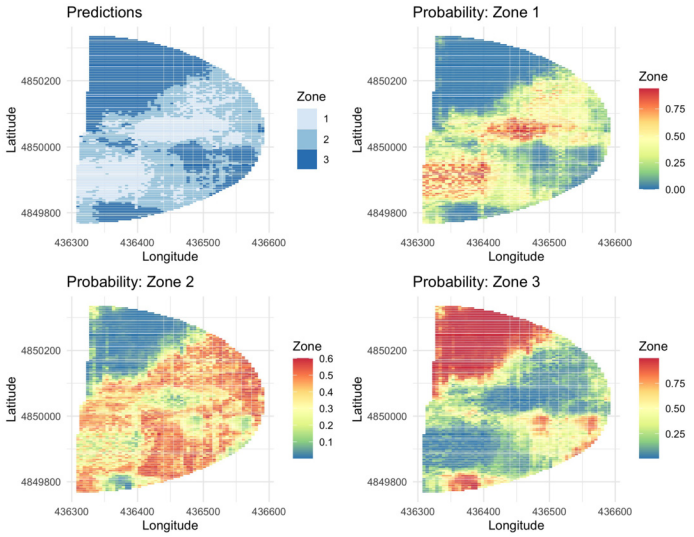


Figure 4: Prediction map for the field as well as the uncertainty in the field. The prediction map is obtained by taking the zone with the highest probability for each location. The uncertainty in the predictions for each location and shown in the probability maps, which show the probability of each location being in each of the three zones.

The spline model and neural network without spatial basis functions had a very similar performance and both seem to be effective approaches to modeling a non-linear relationship. However, this effectiveness was enhanced in both models by the addition of spatial basis functions. The best model was the neural network with the spatial basis functions, which validates the use of the model proposed in this analysis since it accounts for both the non-linear relationships and spatial correlation.

3.3 Field-wide Irrigation Zone Delineation

The primary goal of the analysis was to obtain predictions of the irrigation zones for the whole field for use in the VRI system. The fitted spatial neural network model was used to make predictions of which irrigation zone a location should be classified in across the whole field. Due to Monte Carlo sampling, the uncertainty in the zone delineation was also able to be calculated. Figure 4 shows the highest probability zone by location (top left) along with the probabilities of each location being assigned to each of the three zones. The predictions appear to be fairly spatially contiguous zones (as desired) with the exception of middle of the field in areas between zone 1 and 2.

While zone prediction map in Figure 4 produces zones that are spatially smoothed by the spatial basis function, the covariates produce some non-contiguous regions that would be difficult to implement using variable rate irrigation given the current state of precision agriculture. Rather, for VRI implementation purposes, we desire to further smooth the

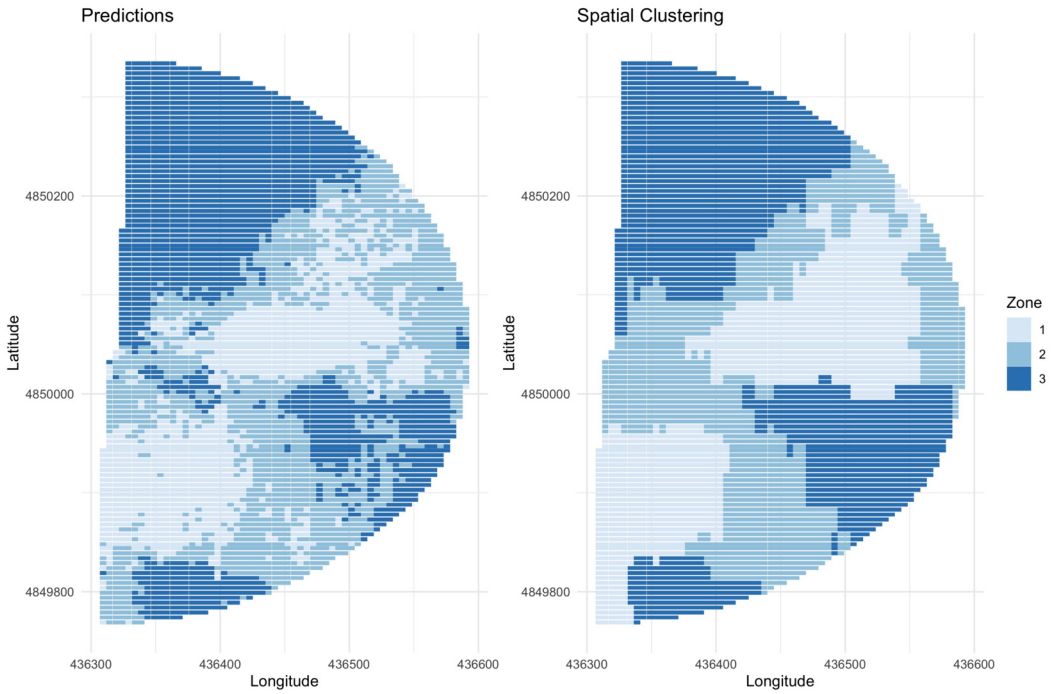


Figure 5: Clustered irrigation zones based on expected zone. The clusters are relatively similar and follow the same general patterns as the predictions but could more reasonably be implemented as irrigation zones using VRI technology.

predicted zones into purely contiguous zones. For purposes of VRI implementation, we use the spatial clustering algorithm of [19] based on finite differences to achieve more contiguous zones. Specifically, we spatially cluster using ward dissimilarity [see 19, for details] based on the expected zone

$$\begin{aligned} E\big(Y(\boldsymbol{s})\big) &= 1 \times \big[\Phi\big(0 - \big(\widehat{f}_L\big(\boldsymbol{X}(\boldsymbol{s})\big) + \widehat{w}(\boldsymbol{s})\big)\big)\big] \\ &\quad + 2 \times \big[\Phi\big(\widehat{c}_2 - \big(\widehat{f}_L\big(\boldsymbol{X}(\boldsymbol{s})\big) + \widehat{w}(\boldsymbol{s})\big)\big) \qquad (3.1) \\ &\quad - \Phi\big(0 - \big(\widehat{f}_L\big(\boldsymbol{X}(\boldsymbol{s})\big) + \widehat{w}(\boldsymbol{s})\big)\big)\big] \\ &\quad + 3 \times \big[1 - \Phi\big(c_2 - \big(\widehat{f}_L\big(\boldsymbol{X}(\boldsymbol{s})\big) + \widehat{w}(\boldsymbol{s})\big)\big)\big] \qquad (3.2) \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative density function. The result from this clustering process is shown in Figure 5 and compared with the original predictions. While there are still some discrepancies between the clustered zones and the original zones, they are relatively similar and the clustering provides contiguous zones that would be viable to be implemented with VRI. We note that clustering the model results (using any spatial clustering method) is not required but is simply beneficial for our application and the implementation of the estimated zones for the growing season.

3.4 Effect of Covariates

While prediction was the primary goal of the analysis, it may also be of interest to examine the effect of the covariates on zone classification. However, one of the drawbacks of using neural network models as we have here is that the parameters have no intuitive interpretation. Indeed, interpretability of neural networks, generally, is an open area of research [29]. One possible solution is to use partial dependence plots (PDPs) and feature importance to intuitively understand the effects of covariates on the response. First, partial dependence plots are used to show the marginal effect of a covariate on the predicted outcome. Mathematically, the partial dependence plot for a covariate, say X_p, is calculated as

$$\text{PDP}(X_p) = \frac{1}{n}\sum_{i=1}^{n}\big(\widehat{f}_L\big(X_p, \boldsymbol{X}_{-p}(\boldsymbol{s}_i)\big) + \widehat{w}(\boldsymbol{s}_i)\big) \qquad (3.3)$$

where $\boldsymbol{X}_{-p}(\boldsymbol{s}_i)$ is the vector of covariates with $X_p(\boldsymbol{s}_i)$ removed and $\widehat{f}_L(\cdot)$ and $\widehat{w}(\cdot)$ is our fitted spatial neural network model. The above is calculated for a grid of X_p in the domain of $X_p(\boldsymbol{s})$ to produce a curve representing the marginal effect of $X_p(\boldsymbol{s})$. Intuitively, the partial dependence measure is calculated by replacing the variable of interest ($X_p(\boldsymbol{s})$) with a fixed singleton value for all observations in the data and averaging the associated prediction across observations. Because we adopted the Bayesian approach, the uncertainty for these partial dependence plots are accounted for as well via Monte Carlo sampling.

Figure 6 shows the partial dependence plots for some of the covariates that were used in the analysis (covariates with no effect were omitted for brevity) where the black line represents the posterior mean of the marginal effect for the

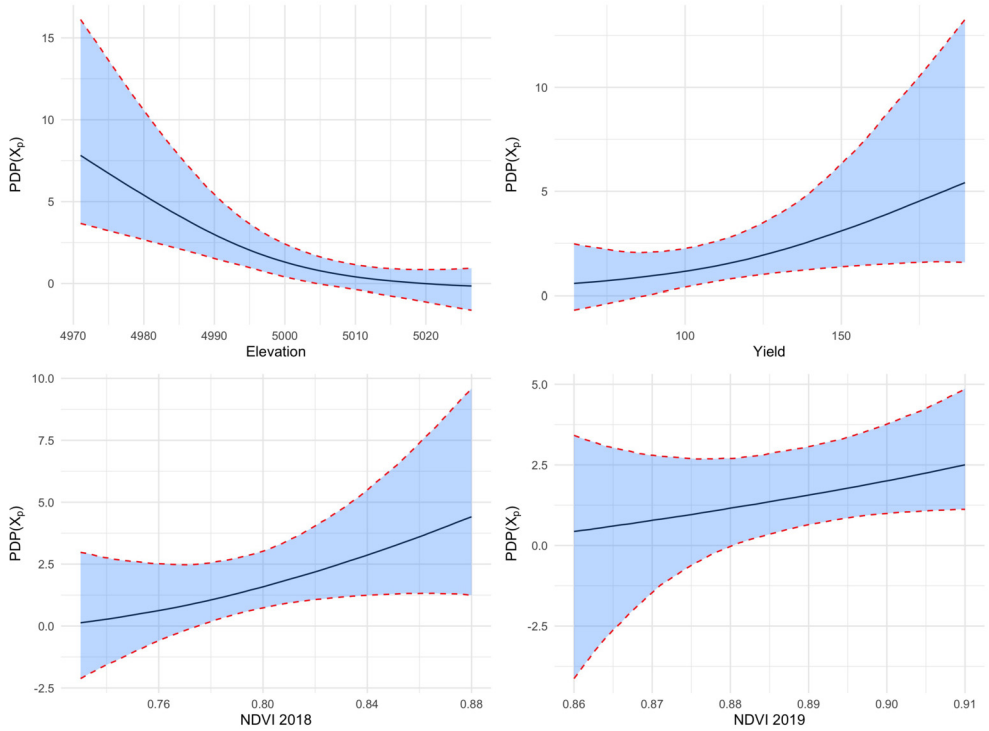


Figure 6: Partial dependence plots for influential covariates. The black line shows the marginal effect for the variable and the shaded area is the 95% credible interval for the marginal effect.

covariate and the shaded region is the 95% credible interval for the marginal effect. Elevation, yield, and the NDVI for 2018 and 2019 seemed to heavily influence the probability of belonging to an irrigation zone. For example, as elevation increases, the probability of belonging to a lower-water zone increases (Zone 1 corresponds to the zone with the smallest VWC which, in turn, would get the most irrigation to compensate). The other covariates appear to have a minimal effect on the predicted zone.

Another common measure for covariate influence on machine learning models is feature importance. For this analysis, the permutation feature importance was used as the feature importance measure. The feature importance was calculated by permuting values of the feature (covariate) and calculating the ARI for the model with the permuted feature. The purpose of permuting the covariate is to break the relationship between the feature and the response. Under permutation, the prediction accuracy should decrease for features that are highly predictive of the response. After calculating the ARI for the model with the permuted feature, the feature importance for the p^{th} covariate $(X_p(\boldsymbol{s}))$ is calculated as:

$$FI_p = ARI_{\text{orig}} - ARI_{\text{perm}} \tag{3.4}$$

where ARI_{orig} is the ARI for the model before permutation and ARI_{perm} is the ARI for the model after feature $X_p(\boldsymbol{s})$ had been permuted. Intuitively, higher FI_p equates to more important features in determining the zone delineation.

The feature importance plot in Figure 7 shows the feature importance for the 10 covariates that were included in the model. In correspondence with the partial dependence plots, the results from the feature importance plot also indicates that elevation, yield, and the NDVI index from 2018 and 2019 are important covariates in determining the irrigation zone. Hence, for other agricultural fields where irrigation zones are desired, collecting the highest importance features will provide the most effective zones for variable rate irrigation.

4. CONCLUSION

This analysis presents a Bayesian spatial neural network model with easily obtainable predictors such as elevation, slope, and past crop yields to be used for irrigation zone delineation. We propose this model as an alternative to the expensive and time consuming process of measuring volumetric water content. The model provides a fusion of statistical modeling and deep learning by harnessing the predictive ability of artificial neural networks, while quantifying uncertainty using Bayesian methods and using spatial modeling to capture spatial correlation in the irrigation zones. The analysis showed that for the field in Rexburg, Idaho, the most influential covariates for delineating irrigation zones were elevation, yield and the NDVI index.

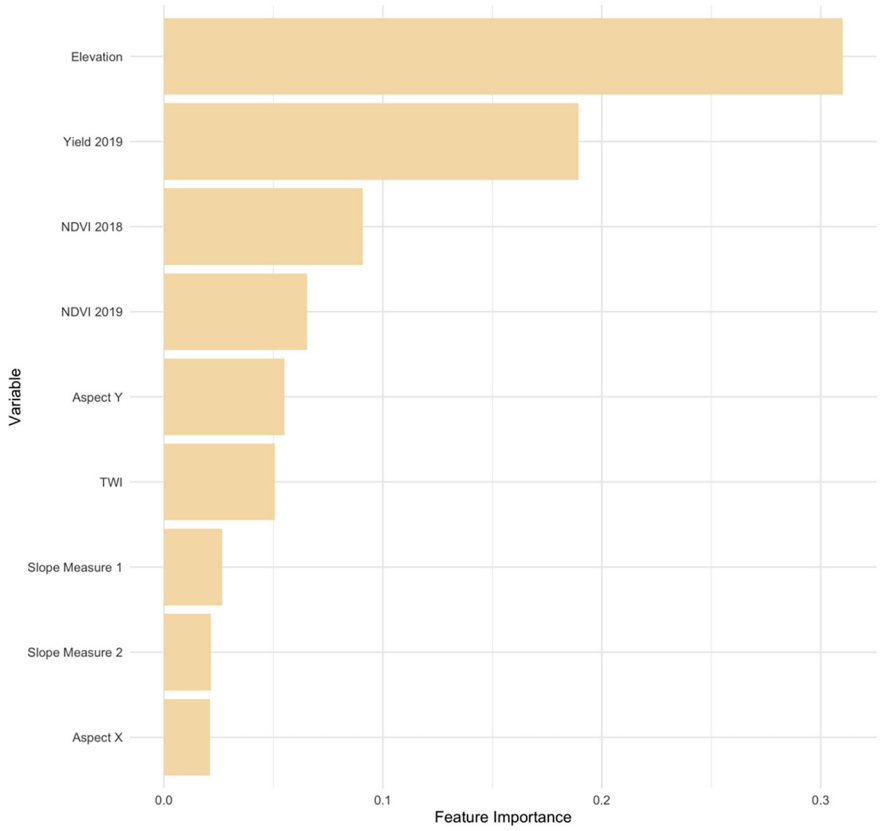


Figure 7: Feature importance for the covariates. Similar to the partial dependence plots, the most important features according to the importance measure is elevation, yield, and the NDVI for 2018 and 2019.

There are some issues with the model presented that could be examined to improve the proposed method. As mentioned, the posterior mixing of the neural network weights was a challenge. This is to be expected given the non-identifiability of neural network weights as discussed in [10]. While we used efforts such as strong prior assumptions (the equivalent of penalization in traditional neural network model fitting), one possible solution to this is to use Hamiltonian Monte Carlo techniques which could integrate the backward-propagation algorithm but in a Bayesian posterior sampling paradigm.

There were 3 irrigation zones chosen for this analysis, but the performance of the model as the number of zones increase has yet to be studied and it is unknown if the model will perform as well when there are more than 3 irrigation zones. Further, perhaps more irrigation zones would increase the efficiency of the VRI technology. Future research could consider the effect of the number of zones on the fitted model.

In addition to the potential issues that should be further studied, there is potential future work that can be done based on the results from the analysis. First, the analysis and results of this model has only been applied to the one field in Rexburg. It would be beneficial to be able to use the model to delineate other fields. The model will be most beneficial when it is portable to other fields and scenarios. At the moment, the efficacy of this model is only known for the field of winter wheat in Rexburg.

In addition, to increase the portability of the model, it would be useful to consider other possible covariates that could be used to determine the volumetric water content. The covariates that were used in the analysis were provided and consequentially there may be other covariates that would be beneficial to delineating irrigation zones. For this analysis, we used ten different covariates and a number of them did not appear to be extremely important in determining the irrigation zone based on the partial dependence and feature importance plots. The impact of adding and subtracting covariates from the model could be examined further.

As mentioned in the Introduction, this research focuses on determining static water management zones. Alternatively, time-varying or dynamic zones could be created to alter the water within a growing season according to the needs of the plant [see 13]. We note that our methods could be used as a foundation for determining dynamic zones but such an application would not only require within-season covariates such as daily precipitation but also require con-

siderations of computational scalability which was less of a concern for this application.

Overall, the use of Bayesian spatial neural network models has the ability to create accurate irrigation zones from easily obtained data about a field without having to put in painstaking effort to determine the volumetric water content. As a result, these models could make the implementation of variable rate irrigation easier for farmers in agricultural fields.

DECLARATIONS

Data Availability. All code and data for this research is available at https://github.com/dteuscher1/Irrigation_NeuralNet.

FUNDING

This work was partially funded by the US-Israel Binational Agricultural Research and Development Fund grant IS-5218-19.

Accepted 21 August 2024

REFERENCES

[1] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. and Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon* **4**(11) 00938.

[2] Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association* **88**(422) 669–679. MR1224394

[3] Bahat, I., Netzer, Y., Ben-Gal, A., Grünzweig, J., Peeters, A. and Cohen, Y. (2019). Comparison of water potential and yield parameters under uniform and variable rate drip irrigation in a cabernet sauvignon vineyard. In *Precision Agriculture'19* 574–577 Wageningen Academic Publishers.

[4] Berrett, C. and Calder, C. A. (2016). Bayesian spatial binary classification. *Spatial Statistics* **16** 72–102. https://doi.org/10.1016/j.spasta.2016.01.004. MR3493089

[5] Bhakta, I., Phadikar, S. and Majumder, K. (2019). State-of-the-art technologies in precision agriculture: a systematic review. *Journal of the Science of Food and Agriculture* **99**(11) 4878–4888.

[6] Bornn, L., Shaddick, G. and Zidek, J. V. (2012). Modeling nonstationary processes through dimension expansion. *Journal of the American Statistical Association* **107**(497) 281–289. https://doi.org/10.1080/01621459.2011.646919. MR2949359

[7] Chen, W., Li, Y., Reich, B. J. and Sun, Y. (2024). Deepkriging: Spatially dependent deep neural networks for spatial prediction. *Statistica Sinica* **34** 291–311. https://doi.org/10.5705/ss.202021.0277. MR4683573

[8] Cisternas, I., Velásquez, I., Caro, A. and Rodríguez, A. (2020). Systematic literature review of implementations of precision agriculture. *Computers and Electronics in Agriculture* **176** 105626.

[9] Cohen, Y., Gogumalla, P., Bahat, I., Netzer, Y., Ben-Gal, A., Lenski, I., Michael, Y. and Helman, D. (2019). Can time series of multispectral satellite images be used to estimate stem water potential in vineyards? In *Precision Agriculture'19* 1–5 Wageningen Academic Publishers.

[10] D'Amour, A., Heller, K., Moldovan, D., Adlam, B., Alipanahi, B., Beutel, A., Chen, C., Deaton, J., Eisenstein, J., Hoffman, M. D. et al. (2020). Underspecification presents challenges for credibility in modern machine learning. *arXiv preprint arXiv:2011.03395*.

[11] Dixon, M. F., Polson, N. G. and Sokolov, V. O. (2019). Deep learning for spatio-temporal modeling: Dynamic traffic flows and high frequency trading. *Applied Stochastic Models in Business and Industry* **35**(3) 788–807. https://doi.org/10.1002/asmb.2399. MR3974250

[12] Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S. and Dehmer, M. (2020). An introductory review of deep learning for prediction models with big data. *Frontiers in Artificial Intelligence* **3** 4.

[13] Fontanet, M., Scudiero, E., Skaggs, T. H., Fernàndez-Garcia, D., Ferrer, F., Rodrigo, G. and Bellvert, J. (2020). Dynamic management zones for irrigation scheduling. *Agricultural Water Management* **238** 106207.

[14] Friedman, J., Hastie, T., Tibshirani, R. et al. (2001) *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York. https://doi.org/10.1007/978-0-387-21606-5. MR1851606

[15] Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *International Conference on Machine Learning* 1050–1059. PMLR.

[16] Ghosh, M., Maiti, T., Kim, D., Chakraborty, S. and Tewari, A. (2004). Hierarchical Bayesian neural networks: an application to a prostate cancer study. *Journal of the American Statistical Association* **99**(467) 601–608. https://doi.org/10.1198/016214504000000665. MR2086386

[17] Haario, H., Saksman, E. and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli* **7**(2) 223–242. https://doi.org/10.2307/3318737. MR1828504

[18] Haghverdi, A., Leib, B. G., Washington-Allen, R. A., Ayers, P. D. and Buschermohle, M. J. (2015). Perspectives on delineating management zones for variable rate irrigation. *Computers and Electronics in Agriculture* **117** 154–167.

[19] Heaton, M. J., Christensen, W. F. and Terres, M. A. (2017). Nonstationary Gaussian process models using spatial hierarchical clustering from finite differences. *Technometrics* **59**(1) 93–101. https://doi.org/10.1080/00401706.2015.1102763. MR3604192

[20] Hedley, C. B. and Yule, I. J. (2009). Soil water status mapping and two variable-rate irrigation scenarios. *Precision Agriculture* **10**(4) 342–355.

[21] Helman, D., Bahat, I., Netzer, Y., Ben-Gal, A., Alchanatis, V., Peeters, A. and Cohen, Y. (2018). Using time series of high-resolution planet satellite images to monitor grapevine stem water potential in commercial vineyards. *Remote Sensing* **10**(10) 1615.

[22] Higgs, M. D. and Hoeting, J. A. (2010). A clipped latent variable model for spatially correlated ordered categorical data. *Computational Statistics & Data Analysis* **54**(8) 1999–2011. https://doi.org/10.1016/j.csda.2010.02.024. MR2640303

[23] Hughes, J. and Haran, M. (2013). Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75**(1) 139–159. https://doi.org/10.1111/j.1467-9868.2012.01041.x. MR3008275

[24] Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A. and Landivar-Bowles, J. (2021). The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Current Opinion in Biotechnology* **70** 15–22.

[25] Katz, L., Naor, A., Litaor, M., Ben-Gal, A., Alchanatis, V., Peres, M., Peeters, A. and Cohen, Y. (2021). Methodology for comparison between uniform and variable rate application in a drip-irrigated peach orchard. In *Precision Agriculture'21* 823–842 Wageningen Academic Publishers.

[26] Khanal, S., Fulton, J. and Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture* **139** 22–32.

[27] Klute, A. (1986). Water retention: laboratory methods. In *Methods of Soil Analysis, Part 1: Physical and Mineralogical Methods* **5** 635–662.

[28] Lee, J., Bahri, Y., Novak, R., Schoenholz, S. S., Pennington, J. and Sohl-Dickstein, J. (2017). Deep neural networks as gaussian processes. *arXiv preprint arXiv:1711.00165*.

[29] Linardatos, P., Papastefanopoulos, V. and Kotsiantis, S. (2020). Explainable AI: A review of machine learning interpretability methods. *Entropy* **23**(1) 18.

[30] Lindblom, J., Lundström, C., Ljung, M. and Jonsson, A. (2017). Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precision Agriculture* **18**(3) 309–331.

[31] Lo, T. H., Heeren, D. M., Mateos, L., Luck, J. D., Martin, D. L., Miller, K. A., Barker, J. B. and Shaver, T. M. (2017). Field characterization of field capacity and root zone available water capacity for variable rate irrigation. *Applied Engineering in Agriculture* **33**(4) 559–572.

[32] Loures, L., Chamizo, A., Ferreira, P., Loures, A., Castanho, R. and Panagopoulos, T. (2020). Assessing the effectiveness of precision agriculture management systems in Mediterranean small farms. *Sustainability* **12**(9) 3765.

[33] Matthews, A. G. d. G., Rowland, M., Hron, J., Turner, R. E. and Ghahramani, Z. (2018). Gaussian process behaviour in wide deep neural networks. *arXiv preprint arXiv:1804.11271*.

[34] McDermott, P. L. and Wikle, C. K. (2019). Bayesian recurrent neural network models for forecasting and quantifying uncertainty in spatial-temporal data. *Entropy* **21**(2) 184. https://doi.org/10.3390/e21020184. MR3923929

[35] Neal, R. M. (2012) *Bayesian Learning for Neural Networks* **118**. Springer Science & Business Media.

[36] Nwankpa, C., Ijomah, W., Gachagan, A. and Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378*.

[37] Ohana-Levi, N., Ben-Gal, A., Peeters, A., Termin, D., Linker, R., Baram, S., Raveh, E. and Paz-Kagan, T. (2021). A comparison between spatial clustering models for determining N-fertilization management zones in orchards. *Precision Agriculture* **22**(1) 99–123.

[38] Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y. and Ben-Gal, A. (2019). A weighted multivariate spatial clustering model to determine irrigation management zones. *Computers and Electronics in Agriculture* **162** 719–731.

[39] O'Shaughnessy, S. A., Evett, S. R., Andrade, A., Workneh, F., Price, J. A. and Rush, C. M. (2015). Site-specific variable rate irrigation as a means to enhance water use efficiency. In *2015 ASABE/IA Irrigation Symposium: Emerging Technologies for Sustainable Irrigation – A Tribute to the Career of Terry Howell, Sr. Conference Proceedings* 1–21. American Society of Agricultural and Biological Engineers.

[40] O'Shaughnessy, S. A., Evett, S. R., Colaizzi, P. D., Andrade, M. A., Marek, T. H., Heeren, D. M., Lamm, F. R. and LaRue, J. L. (2019). Identifying advantages and disadvantages of variable rate irrigation: An updated review. *Applied Engineering in Agriculture* **35**(6) 837–852.

[41] Pang, G., Shen, C., Cao, L. and Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)* **54**(2) 1–38.

[42] Perea, R. G., Daccache, A., Díaz, J. R., Poyato, E. C. and Knox, J. W. (2018). Modelling impacts of precision irrigation on crop yield and in-field water management. *Precision Agriculture* **19**(3) 497–512.

[43] Polson, N. G. and Sokolov, V. (2019). Bayesian regularization: From Tikhonov to horseshoe. *Wiley Interdisciplinary Reviews: Computational Statistics* **11**(4) 1463. https://doi.org/10.1002/wics.1463. MR3999526

[44] Ramachandran, P., Zoph, B. and Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*.

[45] Saha, A., Basu, S. and Datta, A. (2023). Random forests for spatially dependent data. *Journal of the American Statistical Association* **118**(541) 665–683. https://doi.org/10.1080/01621459.2021.1950003. MR4571149

[46] Sauer, A., Gramacy, R. B. and Higdon, D. (2020). Active learning for deep Gaussian process surrogates. *arXiv preprint arXiv:2012.08015*. https://doi.org/10.1080/00401706.2021.2008505. MR4543056

[47] Shand, L. and Li, B. (2017). Modeling nonstationarity in space and time. *Biometrics* **73**(3) 759–768. https://doi.org/10.1111/biom.12656. MR3713110

[48] Sigrist, F. (2022). Gaussian process boosting. *The Journal of Machine Learning Research* **23**(1) 10565–10610. MR4577671

[49] Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y. and Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology* **82**(12) 2635–2670.

[50] Steven, M. and Clark, J. A. (2013) *Applications of Remote Sensing in Agriculture*. Elsevier.

[51] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1) 267–288. MR1379242

[52] Tran, M. -N., Nguyen, N., Nott, D. and Kohn, R. (2020). Bayesian deep net GLM and GLMM. *Journal of Computational and Graphical Statistics* **29**(1) 97–113. https://doi.org/10.1080/10618600.2019.1637747. MR4085867

[53] Wani, M. A., Bhat, F. A., Afzal, S. and Khan, A. I. (2020) *Advances in Deep Learning*. Springer. https://doi.org/10.1007/978-981-13-6794-6. MR3966423

[54] Wikle, C. K. (2019). Comparison of deep neural networks and deep hierarchical models for spatio-temporal data. *Journal of Agricultural, Biological and Environmental Statistics* **24**(2) 175–203. https://doi.org/10.1007/s13253-019-00361-7. MR3945276

[55] Williams, C. K. (1997). Computing with infinite networks. In *Advances in Neural Information Processing Systems* 295–301.

[56] Wójtowicz, M., Wójtowicz, A., Piekarczyk, J. et al. (2016). Application of remote sensing methods in agriculture. *Communications in Biometry and Crop Science* **11**(1) 31–50.

[57] Yari, A., Madramootoo, C. A., Woods, S. A., Adamchuk, V. I. and Huang, H. -H. (2017). Assessment of field spatial and temporal variabilities to delineate site-specific management zones for variable-rate irrigation. *Journal of Irrigation and Drainage Engineering* **143**(9) 04017037.

[58] Zammit-Mangion, A., Kaminski, M. D., Tran, B. -H., Filippone, M. and Cressie, N. (2023). Spatial Bayesian neural networks. *arXiv preprint arXiv:2311.09491*. https://doi.org/10.1016/j.spasta.2024.100825. MR4731204

[59] Zhan, W. and Datta, A. (2023). Neural networks for geospatial data. *arXiv preprint arXiv:2304.09157*.

[60] Zhang, J., Guan, K., Peng, B., Jiang, C., Zhou, W., Yang, Y., Pan, M., Franz, T. E., Heeren, D. M., Rudnick, D. R. et al. (2021). Challenges and opportunities in precision irrigation decision-support systems for center pivots. *Environmental Research Letters* **16**(5) 053003.

[61] Zhao, W., Li, J., Yang, R. and Li, Y. (2017). Crop yield and water productivity responses in management zones for variable-rate irrigation based on available soil water holding capacity. *Transactions of the ASABE* **60**(5) 1659–1667.

Matthew J. Heaton. Department of Statistics, Brigham Young University, 2152 WVB, Provo, UT 84602, USA. E-mail address: mheaton@stat.byu.edu

David Teuscher. Department of Statistics, Brigham Young University, 2152 WVB, Provo, UT 84602, USA. E-mail address: dteuscher.37.12@gmail.com

Neil C. Hansen. Department of Plant & Wildlife Sciences, Brigham Young University, 5108 LSB, Provo, UT 84602, USA. E-mail address: neil_hansen@byu.edu