Machine learning for optimizing complex site-specific management

Yuji Saikai

October 11, 2019

Abstract

Despite the promise of precision agriculture (PA) for increasing the efficiency of crop production by implementing site-specific management, farmers remain skeptical and its utilization rate is lower than expected. A major cause is, unsurprisingly, a lack of concrete approaches to higher profitability. Since PA involves many variables in both controlled management and monitored environment, optimal site-specific management for such high-dimensional cropping systems is considerably more complex than the traditional low-dimensional cases widely studied in the existing literature, calling for a new approach to optimization of site-specific management. We propose an algorithmic approach that enables each farmer to efficiently learn their own site-specific management through on-farm experiments. We demonstrate its performance in simulated environments. Our results show that, for example, the learned site-specific management from 5-year experiments generates $43/ha higher profits with 25 kg/ha less nitrogen fertilizer in one scenario with 150 management variables and $39/ha higher profits with 56 kg/ha less nitrogen fertilizer in another scenario with 864 management variables than the benchmark uniform management. Thus, complex site-specific management can be learned very efficiently, which is indeed more profitable and environmentally sustainable than uniform management.

Keywords: Machine learning, Bayesian optimization, APSIM, precision agriculture, site-specific management, on-farm experiment

1 Introduction

Precision agriculture (PA) holds big promise of modern agriculture in addressing some of the most challenging problems in the 21st century, including farm profitability, food security, and environmental sustainability. All of them requires increase in yield without increase in farm inputs, in other words, higher productivity. Site-specific management emerges as a natural solution to unleashing the potential of PA for higher productivity (Bongiovanni et al., 2004; Cassman, 1999; Gebbers et al., 2010) because, in principle, site-specific management can optimize production at a subfield level, which amounts to finer optimization at a field level.

Despite the big promise, many farmers have perceived PA as big disappointment, causing slow adoption of PA (Bramley, 2009; Schimmelpfennig, 2016). A major reason for the slow adoption is, understandably, a lack of profitability (Castle et al., 2016; Gandorfer et al., 2017). Since every other societal benefit of PA presumes farmers' adoption, convincing farmers is one of the first problems to solve. To be convinced, farmers need to see concrete approaches for implementing site-specific management in their own fields and increasing their profits from the conventional uniform management. However, farmers have received only partial answers to this request from the existing research (Antle, 2019; Antle et al., 2017; Leonard et al., 2017).

The vast majority of site-specific management research investigates only a single management variable, which is often a total amount of nitrogen fertilizer (Anselin et al., 2004; Boyer et al., 2011;

Jin et al., 2017b; Karatay et al., 2019; Thöle et al., 2013). However, multiplicity is the reality in farming where crop yield is an outcome of the complex interactions among a number of management and environmental factors (Bullock et al., 2000; Ruffo et al., 2006). Within a single growing season, a number of decisions are made on types of activities (e.g., tillage, planting, fertilization, spraying, and harvest), their amounts, and their timings. Since the overall profitability is determined by the totality of those decisions, profitability assessment requires more comprehensive analysis than the existing studies. In other words, independent studies of a single management variable provide only partial knowledge and can be misleading as such complex systems typically involve significant nonlinearity (Altieri, 2018; Gliessman, 1990). Thus, we must investigate effects of multiple variables simultaneously.

In addition to the large number of variables in each farm, the multiplicity of farming also implies heterogeneity, that is, a large variety of farms, each of which is unique and operates with different resources (e.g., machinery) in different environments (e.g., soil). Consequently, results from studying "representative" cases are useful for only farmers who face very similar managerial and environmental conditions. This problem is exacerbated in PA where the high dimensionality makes it even more difficult for selected cases to be representative, the problem called "the curse of dimensionality" in mathematics (Bellman, 2015). As mentioned above, since most studies in the current literature only examine low-dimensional scenarios, the representativeness issue is implicitly avoided. But, then, the research community has not genuinely faced up to the very question in the age of PA—how to optimize site-specific management with a large number of variables.

To address the multiplicity problem and facilitate profitability assessment, we propose an algorithmic approach that enables each farmer to efficiently learn their own site-specific management through on-farm experiments. We emphasize the significance of on-farm experimentation (Cook et al., 2013; Griffin, 2018; Piepho et al., 2011), in particular, field-scale experimentation, which is a way to efficiently navigate a high-dimensional variable space and obtain useful information for each individual (Bramley et al., 2011; Panten et al., 2010; Pringle et al., 2004a,b). Our approach is not to construct a rigid empirical model that assumes a specific number and type of variables; instead, it is a machine learning algorithm that is versatile enough to be used for learning optimal site-specific management in a wide range of farming scenarios. First, we mathematically formulate the farmer's problem as profit maximization. Then, we describe our machine learning algorithm to solve it. We demonstrate the algorithm's performance and versatility in two simulated environments with low- and high-complexity. To underline the generality of our approach, we report all the results on a per-hectare basis so that they can be easily scaled. According to our demonstrations, complex site-specific management can be learned very efficiently through on-farm experiments within a few years, which is indeed more profitable and more environmentally sustainable than uniform management.

2 Materials and methods

2.1 Farmer's problem

Imagine a farmer who has access to some PA equipment and is interested in learning site-specific management through on-farm experiments. To conduct field-scale experiments, the farmer divides an entire field into a grid of sites of equal size according to the capacity of the variable rate technologies and yield monitor. That is, in each site, the farmer can uniquely identify and collect a pair of data (x, y)—applying management x and observing the corresponding yield y. The divided field may look as follows:

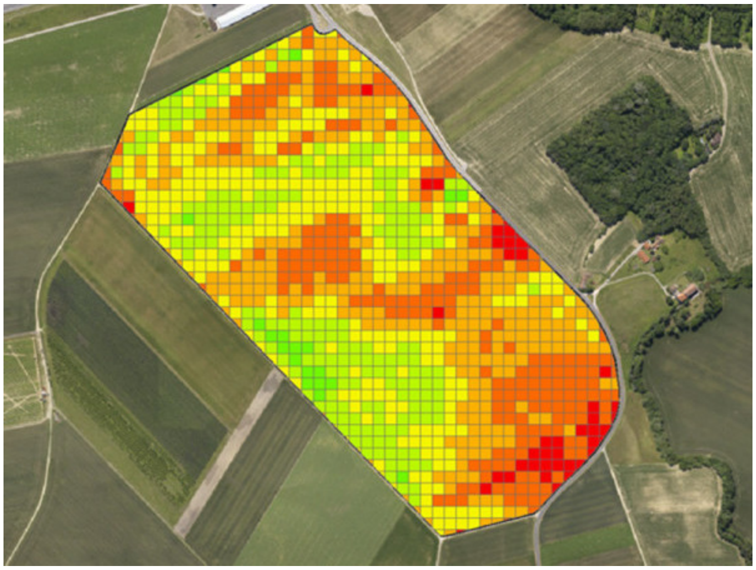


Figure 1: Example crop field divided into a grid of sites

Let M denote the total number of sites. Site $s \in \{1, 2, \dots, M\}$ is characterized by a state variable z_s, to which management x_s is applied. Then, a site-specific profit function is

$$\pi(x_s; z_s) = py(x_s; z_s) - c \cdot x_s,$$

where $y(x_s; z_s)$ is a site-specific yield function, p is an output price, and c is a vector of input prices. Notice that this is technically a partial profit as we subtract only the costs for modeled management x_s. Nonetheless, it is immaterial because our analysis is based on the difference between the profits of the site-specific management and the uniform management.

For conventional low-dimensional yield functions, it is common to use simple concave functions including quadratic (Bachmaier et al., 2009; Meyer-Aurich et al., 2010; Whelan et al., 2012), negative exponential (Edwards et al., 2005; Gaspar et al., 2015), and piecewise linear (Ouedraogo et al., 2018; Park et al., 2018). These simple functions may serve well for answering isolated questions such as optimality of a single management variable under a homogeneous condition. However, in high-dimensional cropping systems, the yield function is a fundamental source of the challenge because its uncertainty increases with the number of variables entering the function $y(\cdot; \cdot)$.

Note that each site need not be recognized as distinct or, equivalently, each z_s need not be distinct. A simple consequence of this assumption is that adjacent sites $\{s_1, s_2, \dots\}$ may have the same value $z_{s_1} = z_{s_2} = \dots$ and form a homogeneous "zone", which receives the same management. This is the framework commonly used to study site-specific management, particularly zone delineation, in the literature (Albornoz et al., 2018; Fraisse et al., 2001; Fridgen et al., 2004; Leroux et al., 2019; Li et al., 2007). Our formulation is more general and contains the existing one as a special case.

Having each site-specific profit defined, a field-level profit is simply the sum of the site-specific profits:

$$\sum_{s=1}^{M} \pi(x_s; z_s) = \sum_{s=1}^{M} py(x_s; z_s) - c \cdot x_s.$$

The farmer's objective is to learn optimal site-specific management x_s^* for all $s \in \{1, 2, \dots, M\}$:

$$(x_1^*, \dots, x_M^*) = \underset{x_1,\dots,x_M}{\operatorname{argmax}} \sum_{s=1}^{M} \pi(x_s; z_s).$$

In contrast, under the uniform management, a single management x is applied to every site s. Therefore, a field-level profit function is:

$$\sum_{s=1}^{M} \pi(x; z_s) = \sum_{s=1}^{M} py(x; z_s) - c \cdot x,$$

and the optimal uniform management $\bar{x}^*$ is:

$$\bar{x}^* = \underset{x}{\operatorname{argmax}} \sum_{s=1}^{M} \pi(x; z_s).$$

2.2 Solution algorithm

We construct an algorithm based on Bayesian optimization (BO) (Brochu et al., 2010; Shahriari et al., 2016), which is a class of numerical optimization techniques used for finding a global optimum of an unknown function. As with many other numerical optimization techniques, BO navigates the search space by examining a point at a time until it locates an acceptable point and halts. Since BO tries to optimize an unknown function, it needs a surrogate model to guide its search. For this purpose, Gaussian process (GP) statistical model is a standard choice in the literature.

BO has two features that makes it suitable for agricultural experiments (Saikai et al., 2019). First, GP as a nonparametric Bayesian model is so flexible that it can adapt to cases in which the objective function takes a complex shape. In high-dimensional precision agriculture, this will likely happen due to strong interactions among many variables involved. Second, BO is in general known for its sample efficiency, which means that BO can locate a good enough point with relatively a small number of examinations. Since agricultural experiments take time before obtaining results, typically a year, the sample efficiency is a desirable feature.

We assume that a single sample is collected from each site, making up M samples in each year. Though depending on the size of a site, in reality, a farmer will likely collect more than one sample from each site. If this is the case, the average value of all samples from the site may be treated as the sample for that site. While the basic BO sequentially processes one sample at a time, we modify it using the "batch expected improvement" acquisition function proposed by Saikai et al. (2019) in order to process M samples at a time. In each year, it proceeds as follows:

1. Prescribes x_s for each s by maximizing the acquisition function $\alpha(x; z_s)$
2. Observes a yield $y(x_s; z_s)$ for all s
3. Computes the corresponding $\pi(x_s; z_s)$ for all s
4. Updates GP with $\{(x_s, z_s, \pi_s)\}_{s=1}^{M}$ and the samples from the preceding years

After completing the planned number of years of experiments, a candidate for x_s^* for each s can be obtained by maximizing the mean function of the learned GP with fixed z_s. Below is the complete algorithm.

Algorithm 1 Batch BO for site-specific management

1: **require:** T, M, $\mathcal{S}$, GP, α

2: **for** $t \in \{1, 2, \ldots, T\}$ **do**

3: $\quad \mathcal{I} \leftarrow \{\ \}$

4: $\quad \widehat{GP} \leftarrow GP$

5: $\quad$ **for** $s \in \{1, 2, \ldots, M\}$ **do**

6: $\quad\quad x_s \leftarrow \operatorname{argmax}_x \alpha(x; z_s)$

7: $\quad\quad \mathcal{I} \leftarrow \mathcal{I} \cup \{(x_s, z_s, \underline{\pi})\}$ where $\underline{\pi} = \min\{\mathcal{S}_\pi\}$

8: $\quad\quad$ Update $\widehat{GP}$ with $\mathcal{S} \cup \mathcal{I}$

9: $\quad$ **for** $s \in \{1, 2, \ldots, M\}$ **do**

10: $\quad\quad y_s \leftarrow Oracle(x_s; z_s)$

11: $\quad\quad \pi_s \leftarrow p y_s - c \cdot x_s$

12: $\quad\quad \mathcal{S} \leftarrow \mathcal{S} \cup \{(x_s, z_s, \pi_s)\}$

13: $\quad$ Update GP with $\mathcal{S}$

14: **return** $M \times T$ number of samples

In terms of notation, T is the total number of years used for experimentation, $\mathcal{S}$ is a set of samples, $\min\{\mathcal{S}_\pi\}$ implies the minimum realized profit, $\alpha(\,\cdot\,; z_s)$ is an acquisition function for site s defined based on GP, and $Oracle(x; z)$ returns an observed yield when x is applied to a site characterized by z. Notice that in Line 2-8 an interim $\widehat{GP}$ is updated with a hypothetical observation ($\underline{\pi}$) so that we can collect a batch of M samples while using the sequential sampling algorithm. As a small detail, in Line 5, site s is chosen in a random order to avoid a systematic bias arising from how we number the sites. Another detail is that, when updating GP, we fit the hyperparameters of GP only to observed data (Line 13) and not to hypothetical data (Line 8).

2.3 Simulation experiments

To construct simulation environments, we make use of the Agricultural Production Systems sIMulator (APSIM), an advanced simulator of cropping systems (Holzworth et al., 2014) being widely used for various purposes such as generating synthetic datasets (Jin et al., 2018, 2019, 2017a; Lobell et al., 2013, 2014, 2015). In each environment, we run the algorithm to learn optimal site-specific management over T years and compare the profit resulting from implementing the learned site-specific management against the benchmark profit resulting from some uniform management. We assume this uniform management to be the one recommended by university extension services.

Depending on specific scenarios, we can customize the algorithm in many ways. Among interesting modifications is how to incorporate observational data that is already collected prior to the experimentation. When comparing against the uniform management as a status quo, a natural dataset assumed is what arises from implementing the uniform management as it represents the existing knowledge. We do so by initializing the GP embedded in the algorithm. Specifically, let $\{(\bar{x}, z_s, \bar{\pi}_s)\}_{s=1}^{M}$ be a set of the uniform management ($\bar{x}$), site characteristics (z_s), and the corresponding profits ($\bar{\pi}_s$). Then, before the algorithm starts a learning process, we fit the GP to these M data points. Note that the use of the uniform management for both benchmark and prior knowledge is merely for simplifying the illustration. In practical applications, farmers may use any benchmark management of interest (either uniform or not) and any existing dataset. Finally, since the algorithm itself involves some randomness, we conduct Monte Carlo simulations and present averaged results over the Monte Carlo samples.

Though the algorithm's applicability is by no means restricted to the scenarios described in this section, we rely on APSIM simulator and construct illustrative test beds within its capability. We simulate a maize production system in Ames, Iowa with the weather data in 2013, which is the most recent year available in APSIM. In terms of management variables, we follow Saikai et al. (2019) and identify six variables $x = (x^1, \dots, x^6)$ in the APSIM maize module:

- x^1 : sowing density (seeds/m^2)
- x^2 : sowing depth (mm)
- x^3 : row spacing (m)
- x^4 : N fertilizer amount before sowing (kg/ha)
- x^5 : N fertilizer amount at sowing (kg/ha)
- x^6 : N fertilizer amount for top dressing (kg/ha)

Using the information from the research and extension services of Iowa State University, we specify the uniform management ($\bar{x}$) as follows:

$$(\bar{x}^1, \bar{x}^2, \bar{x}^3, \bar{x}^4, \bar{x}^5, \bar{x}^6) = (8, 50, 0.76, 67, 67, 67).$$

$\bar{x}^1 = 8$, $\bar{x}^2 = 50$, and $\bar{x}^3 = 0.76$ follow from Elmore (2013) and Farnham (2001). The recommended total nitrogen amount is identified by using the Corn Nitrogen Rate Calculator (Sawyer, 2019), which gives us $\bar{x}^4 + \bar{x}^5 + \bar{x}^6 = 201$. We evenly split it into $\bar{x}^4 = \bar{x}^5 = \bar{x}^6 = 67$. Finally, for calculating profits, the output price is $p = \$0.177$/kg (Duffy, 2013), and the input costs are $c^1 = \$0.00364$/seed and $c^4 = c^5 = c^6 = \$1.29$/kg (Johanns, 2019). We assume no cost for sowing depth and row spacing, which implies the cost vector $c = (c^1, 0, 0, c^4, c^5, c^6)$.

Given the domain knowledge as well as common sense, when the algorithm searches for the optimal management, we restrict the search space to the following:

- $x^1 \in [6.0, 10.0]$ (seeds/m^2)
- $x^2 \in [25, 150]$ (mm)
- $x^3 \in [0.4, 1.0]$ (m)
- $x^4, x^5, x^6 \in [0, 200]$ (kg/ha)

Finally, in our demonstrations, datapoints resulting from the benchmark uniform management $\{(\bar{x}, z_s, \bar{\pi}_s)\}_{s=1}^{M}$ is the only existing dataset incorporated prior to the experimentation. Since they provide no variation in management x, to build up smoothly, in the first year, the algorithm randomly chooses x for each s from the range defined by $\pm 50\%$ of the uniform management.

2.3.1 Scenario A (low complexity)

In this scenarios, we assume that a square field is divided into a grid of 25 sites ($M = 25$). All the sites are distinct, each of which is characterized by a state vector $z_s = (z_s^1, z_s^2)$ where z_s^1 is plant available water capacity (mm) and z_s^2 is organic carbon (%). We set $z_s^1 \in \{231, 259, 288, 317, 346\}$ and $z_s^2 \in \{2.56, 2.88, 3.2, 3.52, 3.84\}$ (± 10 or 20% from the mid values, which are the default values in the APSIM soil module we use).

231, 2.56	259, 2.56	288, 2.56	317, 2.56	346, 2.56
231, 2.88	259, 2.88	288, 2.88	317, 2.88	346, 2.88
231, 3.2	259, 3.2	288, 3.2	317, 3.2	346, 3.2
231, 3.52	259, 3.52	288, 3.52	317, 3.52	346, 3.52
231, 3.84	259, 3.84	288, 3.84	317, 3.84	346, 3.84

Figure 2: Simulated maize field divided into a grid of 25 distinct sites. In each pair of numbers, the first indicates plant available water capacity (mm) and the second indicates organic carbon (%) at that site.

2.3.2 Scenario B (high complexity)

In this scenario, we imagine that the farmer possesses more precise equipment and capable to divide a field into more granular sites, namely, $16 \times 9 = 144$ sites. We also assume that the farmer has conducted more exhaustive soil tests, measuring four state variable (z^1, z^2, z^3, z^4) in each site:

- z^1: plant available water capacity (mm)
- z^2: organic carbon (%)
- z^3: nitrate-N (kg/ha)
- z^4: ammonium-N (kg/ha)

The addition of z^3 and z^4 is because of their significance (Camberato et al., 2017) and availability in APSIM. Notice that nitrogen is also supplied by soil organic matter (z^1) through N-mineralization, creating stronger interactions among management and environmental variables (Sawyer, 2008). Consequently, site-specific management in scenario B is even more complex than in scenario A. We generate a set of state variables for each site in a random but spatially correlated fashion, namely, random walk. Below are summary statistics of the generated state variables for 144 sites.

- z^1: mean = 296, std = 32, min = 199, max = 365
- z^2: mean = 3.19, std = 0.31, min = 2.59, max = 3.90
- z^3: mean = 9.1, std = 0.98, min = 7.0, max = 11.5
- z^4: mean = 10.6, std = 1.6, min = 7.7, max = 14.1

Instead of writing down four numbers at each site, we illustrate the infield variability with a profit map arising from applying the uniform management to the generated field. Since each site receives the same management, the variability in profits indicates the variability in the underlying growing conditions.

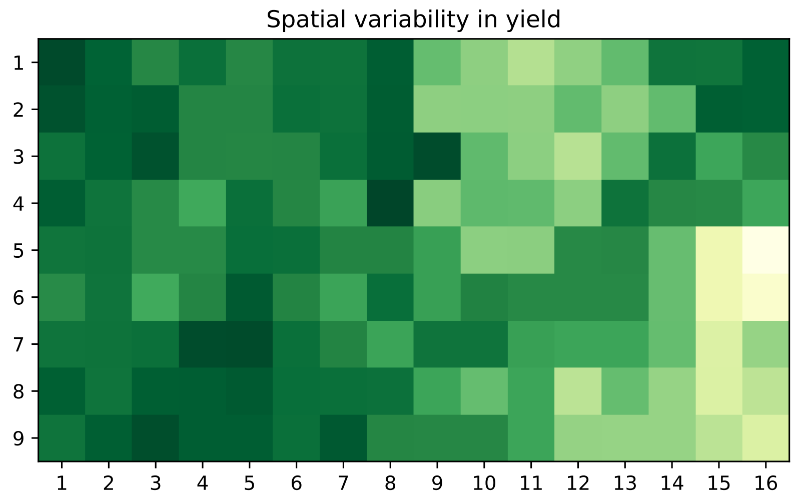


Figure 3: Yield map resulting from the uniform management in scenario B. Axis ticks are added to show the coordinates of each site. Since each site receives the same management, the variability in yield indicates the variability in the underlying growing conditions (state variables).

3 Results

3.1 Scenario A (low complexity)

The following table and figure illustrate field-level profits ($/ha) from the site-specific management learned after T-year experiments. That is, a value for each $T \in \{1, 2, \dots, 10\}$ means a profit if the farmer terminates the experiments after T years and implements the learned site-specific management.

Years (T)	1	2	3	4	5	6	7	8	9	10
Learned	1097	1239	1266	1274	1277	1281	1283	1284	1285	1286
Uniform	1234	1234	1234	1234	1234	1234	1234	1234	1234	1234
Difference	-137	5	32	40	43	47	49	50	51	52
										($/ha)

Table 1: Comparison of field-level profits from implementing the learned and uniform management in scenarios A.

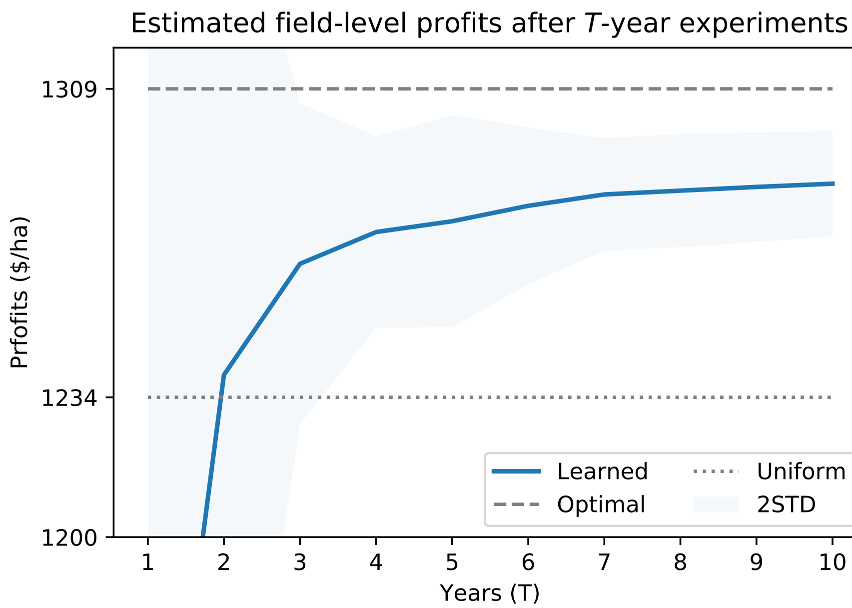


Figure 4: Learning curve of the algorithm with profits plotted against years of experiments in scenario A. The shaded areas indicate two standard deviations around each mean field-level profit over Monte Carlo samples.

The shaded areas indicate two standard deviations around each mean field-level profit over the Monte Carlo samples. The dashed line indicates the profits from the optimal site-specific management, while the dotted line indicates the profits from the uniform management.

Since a field-level profit is the sum of the site-specific profits, next we provide a profit at each site. For the learned profits, we use $T = 5$ as the learning mostly levels off and the deviation from the mean prediction becomes small after four or five years. The following heatmaps illustrate the site-specific profits from the learned and uniform management.

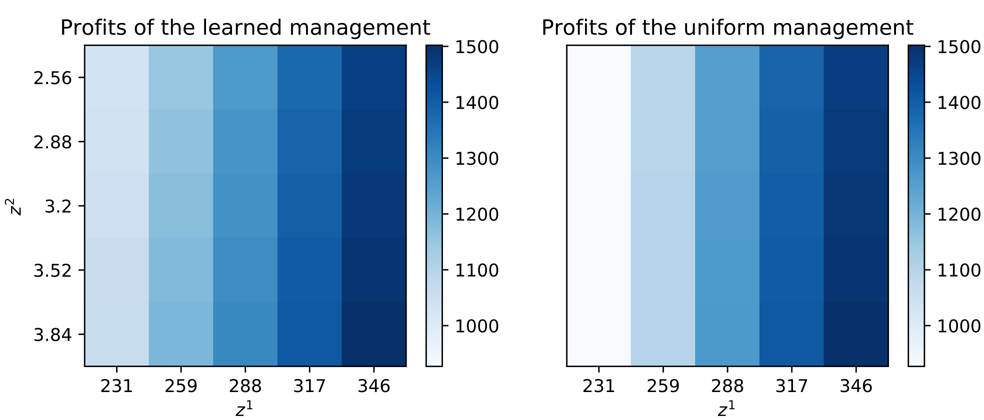


Figure 5: Site-specific profits for the learned and uniform management in scenario A

As indicated in the panel for the uniform management, plant available water capacity (z^1) has much stronger influence on yield/profit than organic carbon (z^2). Also, in both state variables, the higher the values, the more fertile the site is (though hard to see in z^2). Since the overall patterns are quite similar, to highlight the difference, the following figure illustrates the difference at each site.

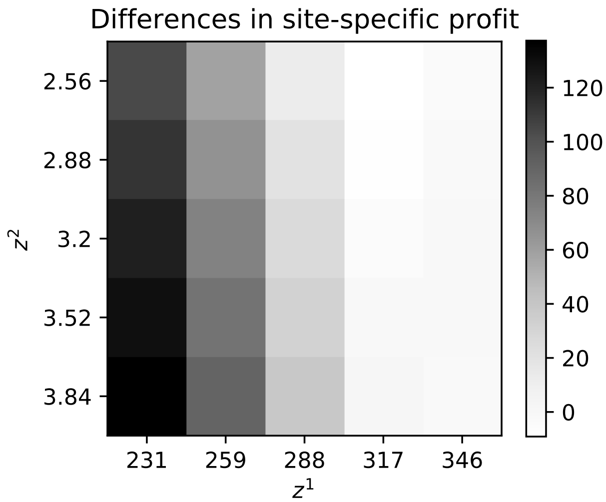


Figure 6: Differences in site-specific profit in scenario A. The maximum difference is \$137/ha at site (231,3.84), whereas the minimum difference is \$-9.1/ha at site (317,2.56).

Finally, we illustrate the learned site-specific management $(x^1, \ldots, x^6)$ after 5-year of experiments.

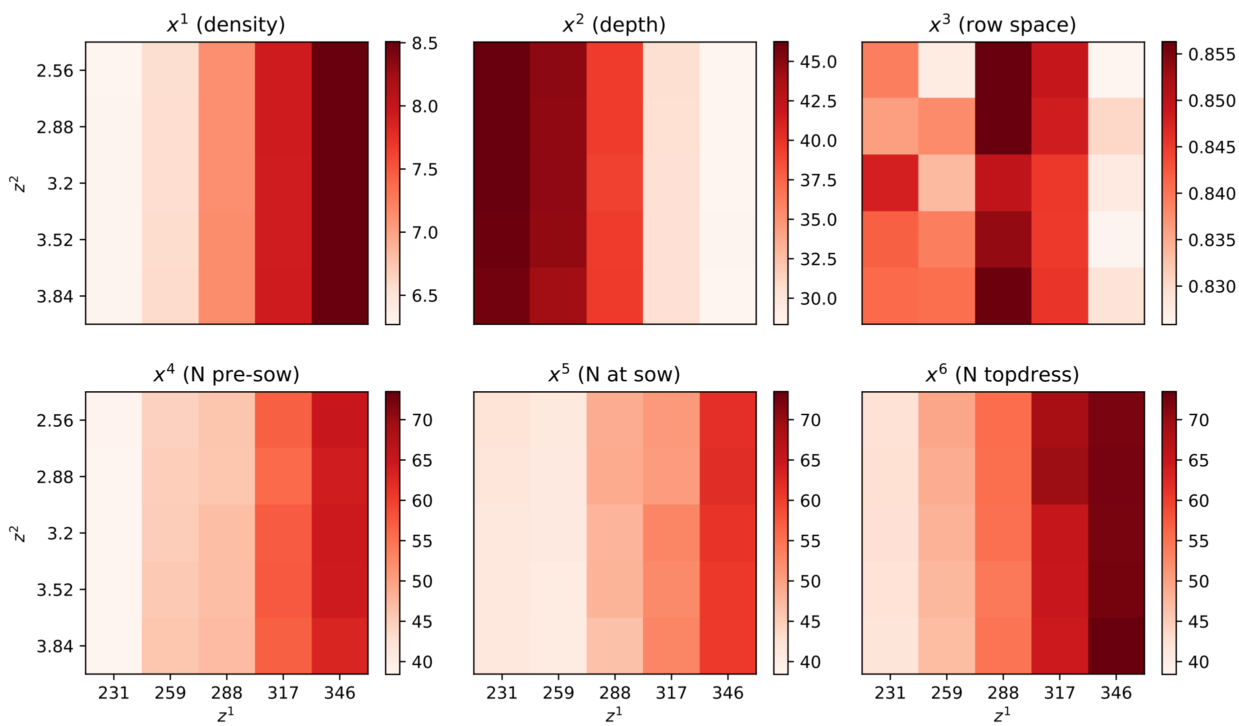


Figure 7: Learned site-specific management after 5-year experiments in scenario A. The average sowing density is 7.3 seeds/m^2, and the average amount of total nitrogen is 156 kg/ha.

The average sowing density is

$$\frac{1}{25}\sum_{s=1}^{25} x_s^1 = 7.3 \text{ seeds/m}^2,$$

and the average amount of total nitrogen fertilizer is

$$\frac{1}{25}\sum_{s=1}^{25}\sum_{i=4}^{6} x_s^i = 156 \text{ kg/ha}.$$

As a result, \$43/ha higher profit is achieved by using 0.7 fewer seeds/m^2 and 45 kg/ha less nitrogen than the uniform management. To further emphasize the generality and robustness of our algorithmic approach, results from other years than 2013 are also provided in Appendices.

3.2 Scenario B (high complexity)

The following table and figure illustrate field-level profits (\$/ha) from the site-specific management learned after T-year experiments. That is, a value for each $T \in \{1, 2, \ldots, 10\}$ means a profit if the farmer terminates the experiments after T years and implements the learned site-specific management.

Years (T)	1	2	3	4	5	6	7	8	9	10
Learned	1320	1322	1329	1331	1334	1335	1333	1335	1337	1338
Uniform	1295	1295	1295	1295	1295	1295	1295	1295	1295	1295
Difference	25	27	34	36	39	40	38	40	42	43
										(\$/ha)

Table 2: Comparison of field-level profits from implementing the learned and uniform management in scenarios B.

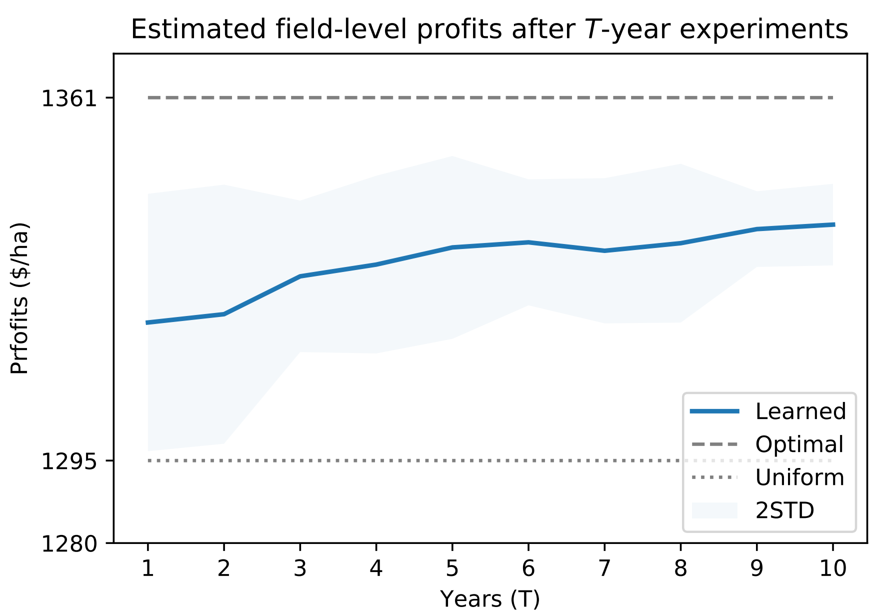


Figure 8: Learning curve of the algorithm with profits plotted against years of experiments in scenario B. The shaded areas indicate two standard deviations around each mean field-level profit over Monte Carlo samples.

The shaded areas indicate two standard deviations around each mean field-level profit over the Monte Carlo samples. The dashed line indicates the profits from the optimal site-specific management, while the dotted line indicates the profits from the uniform management.

The following heatmaps compare the site-specific profits from the learned at $T = 5$ and uniform management.

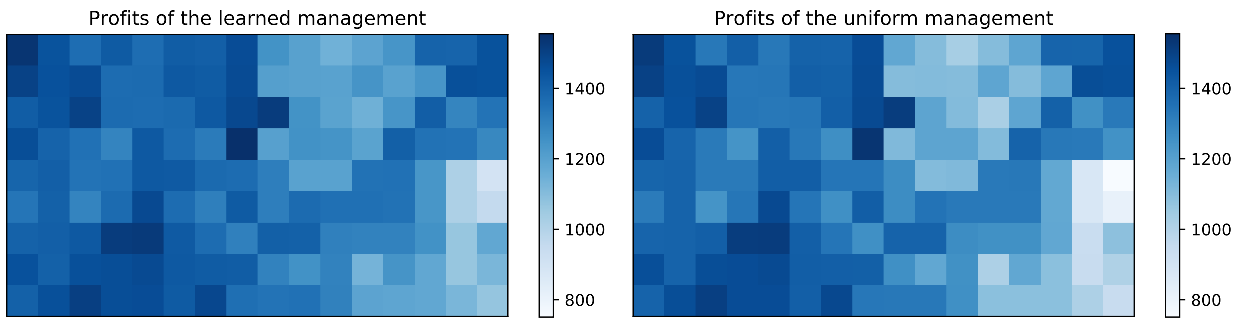


Figure 9: Site-specific profits for the learned and uniform management in scenario B

As in scenario A, the overall patters are very similar. The site-specific management, however, has higher profits (darker colors) in low-yielding sites. The following figure illustrates the difference at each site.

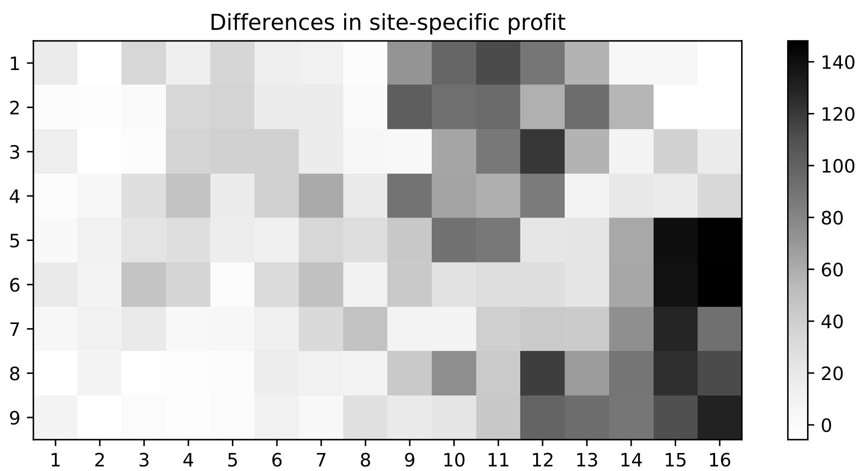


Figure 10: Differences in site-specific profit in scenario B. The maximum difference is \$148/ha at site (16,6), whereas the minimum difference is \$-5.7/ha at site (8,1).

Finally, we illustrate the learned site-specific management $(x^1, \ldots, x^6)$ after 5-year of experiments.

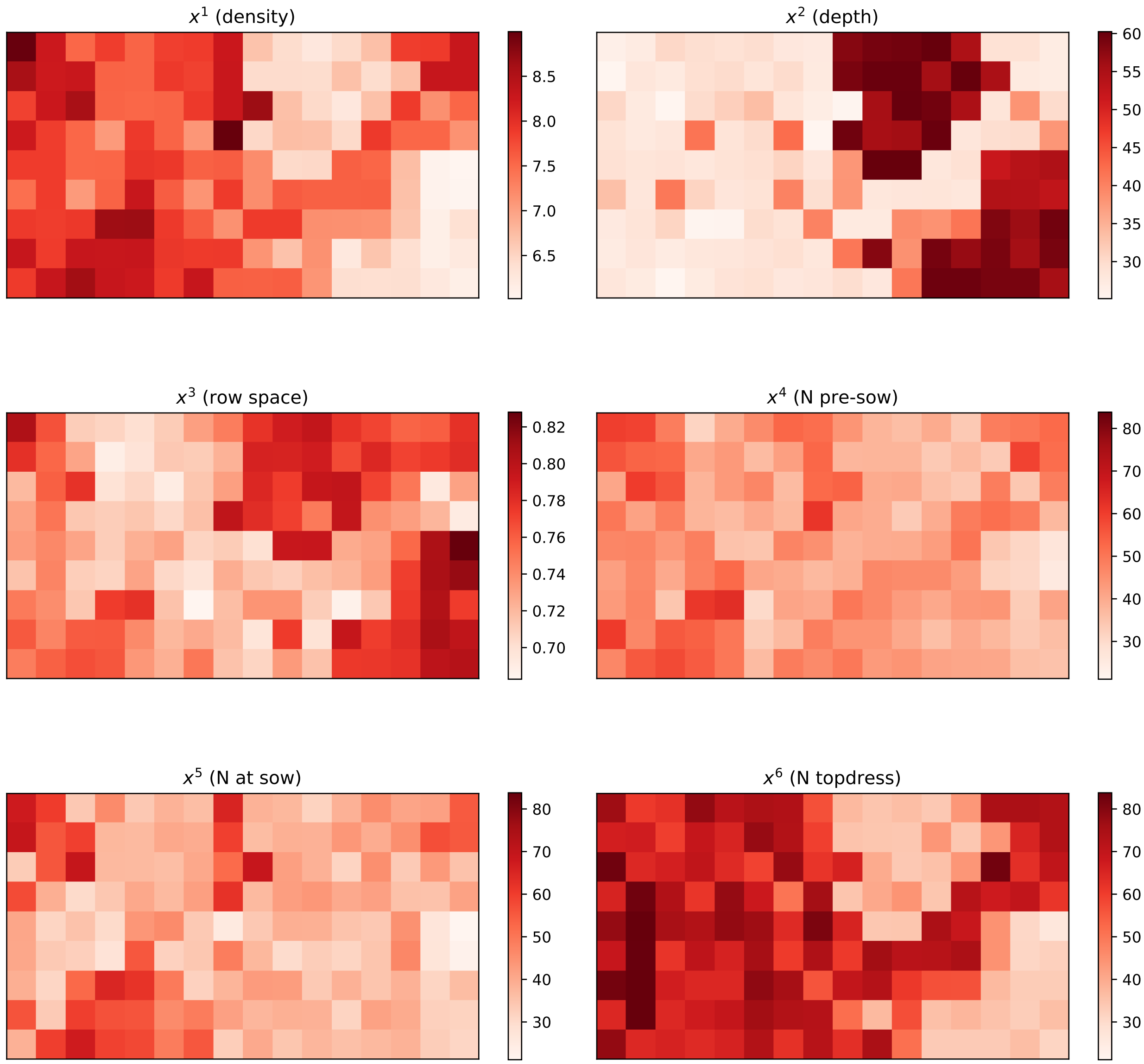


Figure 11: Learned site-specific management after 5-year experiments in scenario B. The average sowing density is 7.4 seeds/m^2, and the average amount of total nitrogen is 145 kg/ha.

The average sowing density is

$$\frac{1}{144}\sum_{s=1}^{144} x_s^1 = 7.4 \text{ seeds/m}^2,$$

and the average amount of total nitrogen fertilizer is

$$\frac{1}{144}\sum_{s=1}^{144}\sum_{i=4}^{6} x_s^i = 145 \text{ kg/ha}.$$

As a result, \$39/ha higher profit is achieved by using 0.6 fewer seeds/m^2 and 56 kg/ha less nitrogen than the uniform management.

4 Discussion

The results for field-level profit in Table 1 and Figure 4 are well aligned with our intuition. When the algorithm starts with little existing data incorporated, it has to explore for good candidate management. As a result, the profit from the leaned management after 1-year experiment is below that of the uniform management. Nonetheless, the algorithm quickly learns and its performance surpasses the performance of the uniform management after 2-year experiments. Thereafter, it continues to learn and widen the performance difference. With 5-year learning, the estimated profit reaches \$1,286 or 98% of the maximum possible profit (\$1,309) in this environment. Given the complexity of the site-specific management, which involves $25 \times 6 = 150$ choice variables, this is considered to be very efficient learning.

As seen in the large shaded area formed by two standard deviations around each mean prediction in Figure 4, due to the random sampling used in the first year, the mean prediction is not very precise during the first few years. However, it dramatically improves after four years, and thereafter its spread continues to shrink. Combined with the increasing mean profits, this is a desirable feature because it implies that no matter how the algorithm starts off, after several years, good site-specific management is consistently learned.

As seen in Figure 5, the higher field-level profit of the learned management is due mainly to its higher profits from low-yielding sites ($z^1 \in \{231, 259\}$). In other words, the uniform management is excessively tailored to the high-yielding conditions ($z^1 \in \{317, 346\}$) leading to the decrease in profitability in the low-yielding sites where it is optimal to put less inputs. As to the overall pattern in the learned site-specific management, we can see in Figure 7 that the higher plant available water capacity (z^1), the more inputs it receives. That is, in the heatmaps for x^1, x^4, x^5, and x^6, the color becomes darker as we move to the right.

We dismiss the patchy looking of x^3 in Figure 7 as an artifact of numerical optimization, which strictly distinguishes two values whenever one results in even a minuscule amount higher than the other. Indeed, the color bar for x^3 has a very small range (0.70 – 0.82 in meter), indicating little practical significance.

Despite \$43/ha higher profit, the 5-year learned management uses only 156 kg/ha nitrogen fertilizer in total, which is much smaller than 201 kg/ha of the uniform management. In terms of the total yield, the learned management produces 9,893 kg/ha, whereas the uniform management produces 10,093 kg/ha. While the management optimization is guided by profit maximization, it turns out to be environmentally more friendly as well. This makes sense because both costs of fertilizer (i.e., to the profitability and to the environment) are aligned and simply the less, the better. However, higher yield does not necessarily coincide with higher profit as yield typically increases with more inputs, which reduces profit.

We have set out to learn site-specific management for maximizing profits because, to realize the other societal benefits of PA (i.e., food security and environmental sustainability), farmers' adoption of PA must precede, which is in turn driven by profitability of PA. So, we have simplified the problem by exclusively focusing on learning profitable management. However, the balance between many benefits of PA need not be this extreme, and it is certainly possible to have a different goal and design an appropriate objective function that guides machine learning in a effective and efficient way.

Finally, despite the promising results, there are several clarifications and limitations to note for real-world implementations and future research.

- We assume no costs for switching management. In reality, however, there will likely be some forms of costs when changing management from site to site. For example, if a management choice is not an amount but a type of fertilizer, it may require human labor and incur non-trivial costs. Also, a high frequency for switching seeding rates may put a excessive strain on and damage an electric motor. For real-world applications, we need to incorporate these costs.

- When variables are theoretically continuous (e.g., seeding rate), we assume that it is possible to choose any arbitrary level and run the algorithm accordingly. In reality, however, choices of continuous variables are constrained for various practical reasons. For example, for a technical reason, we cannot change a fertilizer amount precisely by a fraction of a kilogram. Even though technically possible, it is economically infeasible to change row spacing every time an algorithm recommends a different value. Therefore, for real-world implementations, we need to modify those continuous factors into discrete ones.
- Considering the gap between the learned and the optimal profit (e.g., \$23 for $T = 10$), it is certainly possible to narrow the gap by engineering more sophisticated algorithms such as ones that evolves the search strategies as time goes. We have kept it simple so that we can highlight the generality of our approach and focus on the relative advantage against the uniform management.
- Our approach is limited to one-shot optimization. In other words, a farmer makes all the management decisions at the beginning of the year and wait to see results until the end of the year. In agriculture, many management choices are sequentially made within the year. To handle such more realistic situations, we need dynamic models in which feedback information and learning take place both within and across years.
- As to the temporal variation, we assume a single weather condition and stable site characteristics that prevail over T years. While this may hold or be acceptable in certain circumstances, it will be strong restriction in other circumstances. We are currently developing a model that incorporate temporal variations.
- In addition, we assume out spatial dependency, which means management x_{s_1} for site s_1 affects site s_2, which in turn affects back site s_1, and so on. This may not be negligible under some environments and/or with a very fine grid of sites. Combined with the temporal consideration above, ultimately we need to construct spatiotemporally dynamic models.

5 Conclusions

We have proposed an algorithmic approach to optimizing complex site-specific management with many management and environmental variables. It enables individual farmers to efficiently learn their own site-specific management through on-farm experiments. We have demonstrated its performance using simulated environments. Our results have provided a positive answer to both the learnability of complex site-specific management and the higher profitability than the uniform management. Given the strong support for our approach, real-world field experiments seem warranted to empirically validate the simulated results. We also hope that our approach will empower and incentivize many farmers to utilize their own PA equipment and find more profitable management, which will in turn lead to the greater societal benefits such as food security and environmental sustainability.

References

Albornoz, Enrique M., Alejandra C Kemerer, Romina Galarza, Nicolás Mastaglia, Ricardo Melchiori, and César E Martínez (2018). "Development and evaluation of an automatic software for management zone delineation". In: *Precision Agriculture* 19.3, pp. 463–476.

Altieri, Miguel A (2018). *Agroecology: The science of sustainable agriculture.* 2nd. CRC Press.

Anselin, Luc, Rodolfo Bongiovanni, and Jess Lowenberg-DeBoer (2004). "A spatial econometric approach to the economics of site-specific nitrogen management in corn production". In: *American Journal of Agricultural Economics* 86.3, pp. 675–687.

Antle, John M (2019). "Data, Economics and Computational Agricultural Science". In: *American Journal of Agricultural Economics.*

Antle, John M, James W Jones, and Cynthia Rosenzweig (2017). "Next generation agricultural system models and knowledge products: Synthesis and strategy". In: *Agricultural Systems* 155, pp. 179–185.

Bachmaier, Martin and Markus Gandorfer (2009). "A conceptual framework for judging the precision agriculture hypothesis with regard to site-specific nitrogen application". In: *Precision agriculture* 10.2, p. 95.

Bellman, Richard E (2015). *Adaptive control processes: a guided tour.* Vol. 2045. Princeton university press.

Bongiovanni, Rodolfo and Jess Lowenberg-DeBoer (2004). "Precision agriculture and sustainability". In: *Precision agriculture* 5.4, pp. 359–387.

Boyer, Christopher N., B. Wade Brorsen, John B. Solie, and William R. Raun (2011). "Profitability of variable rate nitrogen application in wheat production". In: *Precision Agriculture* 12.4, pp. 473–487.

Bramley, RGV (2009). "Lessons from nearly 20 years of Precision Agriculture research, development, and adoption as a guide to its appropriate application". In: *Crop and Pasture Science* 60.3, pp. 197–217.

Bramley, RGV, KJ Evans, KJ Dunne, and DL Gobbett (2011). "Spatial variation in response to 'reduced input'spray programs for powdery mildew and botrytis identified through whole-of-block experimentation". In: *Australian Journal of Grape and Wine Research* 17.3, pp. 341–350.

Brochu, Eric, Vlad M Cora, and Nando De Freitas (2010). "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning". In: *arXiv preprint arXiv:1012.2599.*

Bullock, David S and Donald G Bullock (2000). "From agronomic research to farm management guidelines: A primer on the economics of information and precision technology". In: *Precision Agriculture* 2.1, pp. 71–101.

Camberato, Jim and R.L. Nielsen (2017). *Soil Sampling to Assess Current Soil N Availability.* Purdue University. URL: https://www.agry.purdue.edu/ext/corn/news/timeless/assessavailablen.html.

Cassman, K. G. (1999). "Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture". In: *Proceedings of the National Academy of Sciences* 96.11, pp. 5952–5959.

Castle, Michael H, Bradley D Lubben, and Joe D Luck (2016). *Factors Influencing Producer Propensity for Data Sharing & Opinions Regarding Precision Agriculture and Big Farm Data.* UNL Digital Commons.

Cook, Simon, James Cock, Thomas Oberthür, and Myles Fisher (2013). "On-farm experimentation". In: *Better Crops* 97.4, pp. 17–20.

Duffy, Michael (2013). *Estimated Costs of Crop Production in Iowa - 2013*. Estimated Costs of Crop Production in Iowa — Ag Decision Maker. URL: http://econ2.econ.iastate.edu/faculty/duffy/documents/EstimatedCostsofCropProduction2013.pdf.

Edwards, Jeffrey T and Larry C Purcell (2005). "Soybean yield and biomass responses to increasing plant population among diverse maturity groups". In: *Crop Science* 45.5, pp. 1770–1777.

Elmore, Roger W (2013). *Corn Planting FAQs — Integrated Crop Management*. URL: https://crops.extension.iastate.edu/cropnews/2013/04/corn-planting-faqs (visited on 09/28/2019).

Farnham, Dale (2001). *Corn Planting Guide*.

Fraisse, CW, KA Sudduth, and NR Kitchen (2001). "Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity". In: *Transactions of the ASAE* 44.1, p. 155.

Fridgen, Jon J, Newell R Kitchen, Kenneth A Sudduth, Scott T Drummond, William J Wiebold, and Clyde W Fraisse (2004). "Management Zone Analyst (MZA): Software for Subfield Management Zone Delineation". In: *Agronomy Journal* 96, pp. 100–108.

Gandorfer, Markus and Andreas Meyer-Aurich (2017). "Economic Potential of Site-Specific Fertiliser Application and Harvest Management". In: *Precision Agriculture: Technology and Economic Perspectives*. Ed. by Søren Marcus Pedersen and Kim Martin Lind. Cham: Springer International Publishing, pp. 79–92.

Gaspar, Adam P, Paul D Mitchell, and Shawn P Conley (2015). "Economic risk and profitability of soybean fungicide and insecticide seed treatments at reduced seeding rates". In: *Crop Science* 55.2, pp. 924–933.

Gebbers, Robin and Viacheslav I Adamchuk (2010). "Precision agriculture and food security". In: *Science* 327.5967, pp. 828–831.

Gliessman, Stephen R. (1990). "Agroecology: Researching the Ecological Basis for Sustainable Agriculture". In: *Agroecology: Researching the Ecological Basis for Sustainable Agriculture*. Ed. by Stephen R. Gliessman. New York, NY: Springer New York, pp. 3–10.

Griffin, Terry (2018). *Collating and analysing small data to make big decisions – Can it improve farm productivity and profitability?* Grain Research and Development Corporation. URL: https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2018/02/collating-and-analysing-small-data-to-make-big-decisions.

Holzworth, Dean P, Neil I Huth, Peter G deVoil, Eric J Zurcher, Neville I Herrmann, Greg McLean, Karine Chenu, Erik J van Oosterom, Val Snow, and Chris Murphy (2014). "APSIM–evolution towards a new generation of agricultural systems simulation". In: *Environmental Modelling & Software* 62, pp. 327–350.

Jin, Zhenong, Elizabeth A Ainsworth, Andrew D B Leakey, and David B Lobell (2018). "Increasing drought and diminishing benefits of elevated carbon dioxide for soybean yields across the US Midwest". In: *Global Change Biology* 24.2, e522–e533.

Jin, Zhenong, Sotirios V Archontoulis, and David B Lobell (2019). "How much will precision nitrogen management pay off? An evaluation based on simulating thousands of corn fields over the US Corn-Belt". In: *Field Crops Research* 240, pp. 12–22.

Jin, Zhenong, George Azzari, and David B Lobell (2017a). "Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches". In: *Agricultural and Forest Meteorology* 247, pp. 207–220.

Jin, Zhenong, Rishi Prasad, John Shriver, and Qianlai Zhuang (2017b). "Crop model- and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system". In: *Precision Agriculture* 18.5, pp. 779–800.

Johanns, Ann (2019). *Iowa cash corn and soybean prices*. Cash Corn and Soybean Prices — Ag Decision Maker. URL: https://www.extension.iastate.edu/agdm/crops/html/a2-11.html.

Jones, Donald R, Matthias Schonlau, and William J Welch (1998). "Efficient global optimization of expensive black-box functions". In: *Journal of Global optimization* 13.4, pp. 455–492.

Karatay, Yusuf Nadi and Andreas Meyer-Aurich (2019). "Profitability and downside risk implications of site-specific nitrogen management with respect to wheat grain quality". In: *Precision Agriculture*.

Leonard, Emma, Rohan Rainbow, A Laurie, David Lamb, R Llewellyn, Ed Perrett, Jay Sanderson, Andrew Skinner, T Stollery, and Leanne Wiseman (2017). "Accelerating precision agriculture to decision agriculture: Enabling digital agriculture in Australia". In:

Leroux, Corentin and Bruno Tisseyre (2019). "How to measure and report within-field variability: a review of common indicators and their sensitivity". In: *Precision Agriculture* 20.3, pp. 562–590.

Li, Yan, Zhou Shi, Feng Li, and Hong-Yi Li (2007). "Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land". In: *Computers and Electronics in Agriculture* 56.2, pp. 174–186.

Lobell, David B, Graeme L Hammer, Greg McLean, Carlos Messina, Michael J Roberts, and Wolfram Schlenker (2013). "The critical role of extreme heat for maize production in the United States". In: *Nature Climate Change* 3.5, p. 497.

Lobell, David B, Michael J Roberts, Wolfram Schlenker, Noah Braun, Bertis B Little, Roderick M Rejesus, and Graeme L Hammer (2014). "Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest". In: *Science* 344.6183, p. 516.

Lobell, David B, David Thau, Christopher Seifert, Eric Engle, and Bertis Little (2015). "A scalable satellite-based crop yield mapper". In: *Remote Sensing of Environment* 164, pp. 324–333.

Meyer-Aurich, Andreas, Alfons Weersink, Markus Gandorfer, and Peter Wagner (2010). "Optimal site-specific fertilization and harvesting strategies with respect to crop yield and quality response to nitrogen". In: *Agricultural Systems* 103.7, pp. 478–485.

Moćkus, J, V Tiesis, and A Źilinskas (1978). "The Application of Bayesian Methods for Seeking the Extremum. Vol. 2". In: Dixon, L and G Szego. *Toward Global Optimization*. Vol. 2. Amsterdam, The Netherlands: Elsevier.

Ouedraogo, Frederic and B Wade Brorsen (2018). "Hierarchical Bayesian Estimation of a Stochastic Plateau Response Function: Determining Optimal Levels of Nitrogen Fertilization". In: *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 66.1, pp. 87–102.

Panten, K., R. G. V. Bramley, R. M. Lark, and T. F. A. Bishop (2010). "Enhancing the value of field experimentation through whole-of-block designs". In: *Precision Agriculture* 11.2, pp. 198–213.

Park, Eunchun, Wade Brorsen, and Xiaofei Li (2018). *How to Use Yield Monitor Data to Determine Nitrogen Recommendations: Bayesian Kriging for Location Specific Parameter Estimates*. Agricultural and Applied Economics Association.

Piepho, Hans-Peter, Christel Richter, Joachim Spilke, Karin Hartung, Arndt Kunick, and Heinrich Thöle (2011). "Statistical aspects of on-farm experimentation". In: *Crop and Pasture Science* 62.9, pp. 721–735.

Pringle, M. J., S. E. Cook, and A. B. McBratney (2004a). "Field-Scale Experiments for Site-Specific Crop Management. Part I: Design Considerations". In: *Precision Agriculture* 5.6, pp. 617–624.

Pringle, M. J., A. B. McBratney, and S. E. Cook (2004b). "Field-Scale Experiments for Site-Specific Crop Management. Part II: A Geostatistical Analysis". In: *Precision Agriculture* 5.6, pp. 625–645.

Rasmussen, Carl Edward and Christopher K Williams (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Ruffo, Matías L, Germán A Bollero, David S Bullock, and Donald G Bullock (2006). "Site-specific production functions for variable rate corn nitrogen fertilization". In: *Precision Agriculture* 7.5, pp. 327–342.

Saikai, Yuji, Vivak Patel, Lucía Gutiérrez, Brian Luck, Jed Colquhoun, Shawn P Conley, and Paul D Mitchell (2019). "Adaptive experimental design using Bayesian optimization to improve the cost efficiency of small-plot field trials".

Sawyer, John (2008). *Measuring the Nitrogen Status — Integrated Crop Management*. Iowa State University. URL: https://crops.extension.iastate.edu/cropnews/2008/06/measuring-nitrogen-status.

— (2019). *The Corn Nitrogen Rate Calculator*. URL: http://cnrc.agron.iastate.edu/nRate.aspx (visited on 09/28/2019).

Schimmelpfennig, David (2016). *Farm profits and adoption of precision agriculture*. United States Department of Agriculture, Economic Research Service.

Shahriari, Bobak, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas (2016). "Taking the human out of the loop: A review of bayesian optimization". In: *Proceedings of the IEEE* 104.1, pp. 148–175.

Snoek, Jasper, Hugo Larochelle, and Ryan P Adams (2012). "Practical Bayesian optimization of machine learning algorithms". In: Advances in neural information processing systems, pp. 2951–2959.

Stein, Michael L (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Science & Business Media.

Thöle, Heinrich, Christel Richter, and Detlef Ehlert (2013). "Strategy of statistical model selection for precision farming on-farm experiments". In: *Precision Agriculture* 14.4, pp. 434–449.

Whelan, B M, J A Taylor, and A B McBratney (2012). "A small strip approach to empirically determining management class yield response functions and calculating the potential financial net wastage associated with whole-field uniform-rate fertiliser application". In: *Field Crops Research* 139, pp. 47–56.

Appendices

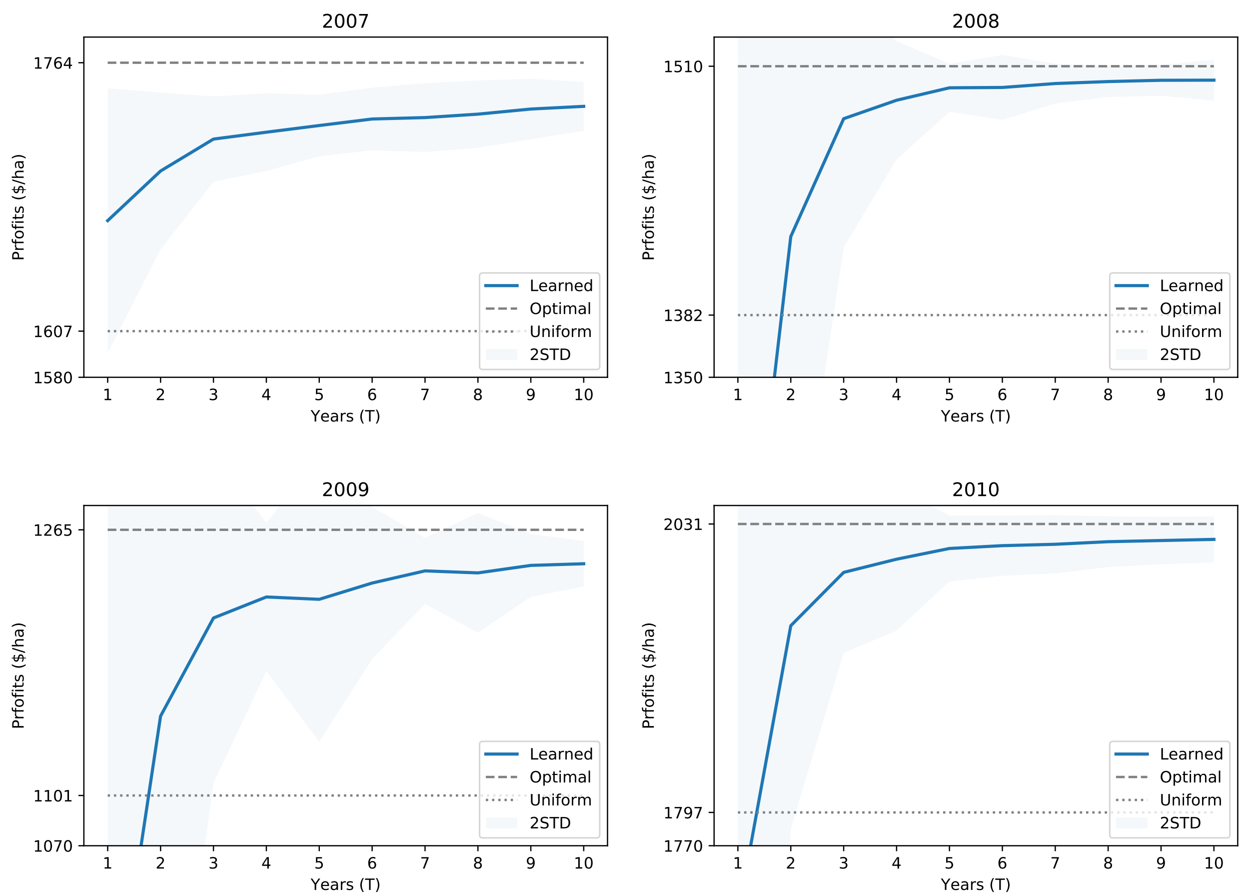
APSIM configuration

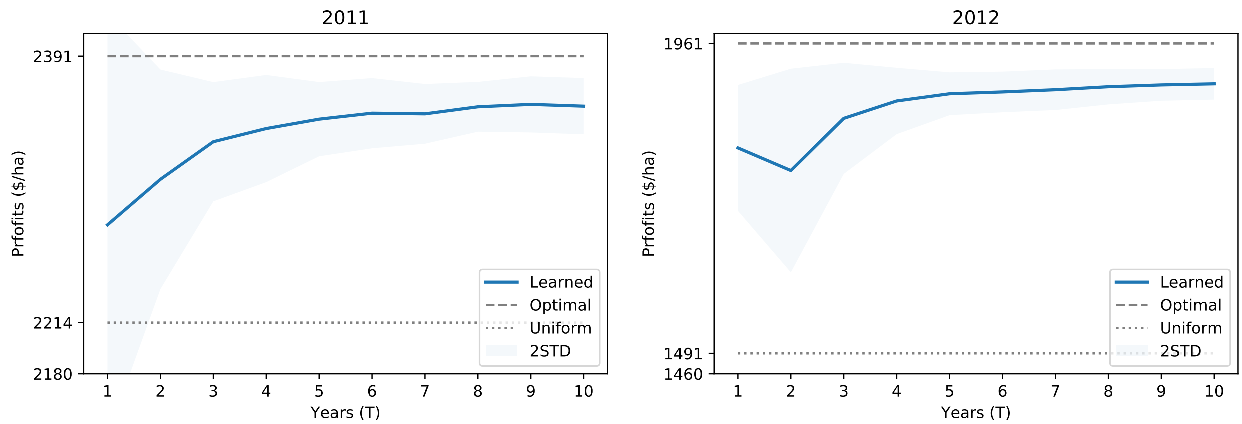
As a basis, we use the Continuous Maize module in APSIM. Then, we modify its default settings as follows. To simulate maize production in Ames, Iowa.

- Metfile: USA_Iowa_Ames.met
- Calendar: Jan 1, 2013 - Dec 31, 2013
- Cultivar: Pioneer 3394
- Sowing window START data: 15-apr
- Sowing window END data: 2-may
- Soil: Iowa Nicollet soil series
- Initial nitrogen: 0 kg/ha for both NO3 and NH4 for scenario A
- Initial water: 80% filled from top

Sensitivity analysis

In addition to two simulated environments with low- and high-complexity, to further emphasize the generality and robustness of our algorithmic approach, we conducted simulation experiments in different years than 2013. Since output price, input prices, and weather are all dependent on a particular year, the differences in year provide different environments for profit maximization. The price information for each year was obtained from the same sources (Duffy, 2013; Johanns, 2019). Note that all sensitivity analysis was conducted under the environments with low complexity, because of the significantly greater computational resources required in environments with high complexity. While there were considerable variations in both the growing and economic conditions across the different years, overall, the algorithm is quite versatile and able to learn good site-specific management within a few years in every environment. Similar to Figure 4, for each environment, we plot estimated field-level profits after T-year experiments.





Gaussian process

Gaussian process is a Bayesian nonparametric model, and its behavior is largely dependent on a choice of kernel and its hyperparameters (Rasmussen et al., 2006). A kernel is a function that returns a similarity measure $k(x, x')$ between two points x and x'. We use the Matérn kernel—a popular class of isotropic stationary kernels.

$$k_\nu(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu B_\nu\left(\sqrt{2\nu}\frac{d}{\rho}\right),$$

where Γ is the gamma function, B_ν is the modified Bessel function of the second kind, and d is a metric often induced by the Euclidean norm, i.e. $d = \|x - x'\|$. The Matérn kernel is characterized by two hyperparameters ν and ρ, which control, respectively, the smoothness and the scaling of distance. As standard in applied work, we do not estimate but rather handpick ν and write as Matérn$_\nu$ or $k_\nu(x, x')$. To simplify the notation, let r denote the scaled distance, $r = d/\rho$. An important property of the Matérn kernel is that when $\nu = p + 1/2, p \in \mathbb{N}$, it can be written as a product of an exponential and a polynomial of order p:

$$k_{p+1/2}(x, x') = \sigma^2 \exp\left(-\sqrt{2p+1}r\right) \frac{p!}{(2p)!} \sum_{i=0}^{p} \frac{(p+i)!}{i!(p-i)!}(2\sqrt{2p+1}r)^{p-i}.$$

Common choices of ν are $1/2, 3/2, 5/2$ and ∞, with each of which the kernel reduces to, respectively,

$$\begin{aligned}
k_{1/2}(x, x') &= \sigma^2 \exp(-r) \\
k_{3/2}(x, x') &= \sigma^2 \exp\left(-\sqrt{3}r\right)\left(1 + \sqrt{3}r\right) \\
k_{5/2}(x, x') &= \sigma^2 \exp\left(-\sqrt{5}r\right)\left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) \\
k_\infty(x, x') &= \lim_{\nu\to\infty} k_\nu(x, x') = \sigma^2 \exp\left(-\frac{1}{2}r^2\right).
\end{aligned}$$

Matérn$_\infty$ is also known as squared exponential kernel or radial basis function. Following Snoek et al. (2012) and Stein (1999), we avoid squared exponential and use Matérn with $\nu = 3/2$ for our algorithm. The following figure plots $k_\nu(x, x')$ with $\sigma^2 = \rho = 1$ for $\nu \in \{1/2, 3/2, \infty\}$.

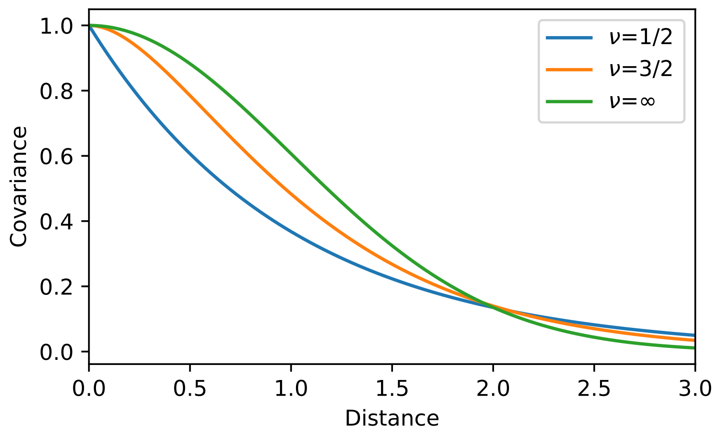


Figure 12: Matérn$_\nu$ kernels

Expected improvement acquisition function

In Bayesian optimization, an algorithm prescribes the next sampling point x based on how we value the mean and variance at x estimated by the accompanying GP. Specifically, the recommendation x_t for the next round t is determined by maximizing an acquisition function $\alpha(x|D_{t-1})$:

$$x_t = \operatorname*{argmax}_x \alpha(x|D_{t-1}),$$

where D_{t-1} is the data used to fit the GP at round $t-1$. The acquisition function is a reflection of the underlying utility of the next sample or our preference in selecting the next sampling point. It is heuristic and designed to trade off exploration of the search space and exploitation of the current promising areas. There are a number of acquisitions functions proposed in the literature. One of the popular acquisition functions is called expected improvement, which is constructed based on the following intuitive idea. Let y^* be the maximum value observed up until round $t-1$, i.e. $y^* = \max\{y_1, \dots, y_{t-1}\}$. Then, we may define "improvement" at point x at round t to be

$$\max\{0, GP(x) - y^*\},$$

which is random as $GP(x)$ is a random function. Thus, the expected improvement acquisition function is defined to be:

$$\alpha_{EI}(x|D_{t-1}) = \mathbb{E}[\max\{0, GP(x) - y^*\}|D_{t-1}].$$

When using Gaussian process, at each point x in the domain, we have $GP(x) \sim \mathcal{N}(\mu(x), \sigma(x))$, which allows the expected improvement to have a closed form (Jones et al., 1998; Moćkus et al., 1978):

$$\alpha_{EI}(x|D_{t-1}) = \begin{cases} (\mu(x) - y^*)\Phi\left(\frac{\mu(x)-y^*}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{\mu(x)-y^*}{\sigma(x)}\right) & \text{if } \sigma(x) > 0 \\ 0 & \text{if } \sigma(x) = 0 \end{cases},$$

where Φ is the standard normal cumulative distribution function and ϕ is the standard normal probability density function.