Machine learning for optimizing complex site-specific management

Yuji Saikai*[1], Vivak Patel[2] and Paul D. Mitchell[1]

[1]Department of Agricultural and Applied Economics, University of Wisconsin–Madison
[2]Department of Statistics, University of Wisconsin–Madison

Abstract

Despite the promise of precision agriculture for increasing the productivity by implementing site-specific management, farmers remain skeptical and its utilization rate is lower than expected. A major cause is a lack of concrete approaches to higher profitability. When involving 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 paradigm shift in optimization of site-specific management. We develop a machine learning algorithm that enables farmers to efficiently learn their own site-specific management through on-farm experiments. We test its performance in two simulated scenarios—one of medium complexity with 150 management variables and one of high complexity with 864 management variables. Results show that, relative to uniform management, site-specific management learned from 5-year experiments generates $43/ha higher profits with 25 kg/ha less nitrogen fertilizer in the first scenario and $40/ha higher profits with 55 kg/ha less nitrogen fertilizer in the second scenario. Thus, complex site-specific management can be learned efficiently and be more profitable and environmentally sustainable than uniform management.

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

1 Introduction

Modern agriculture faces some of the most challenging problems of the 21st century, including food security, farm profitability and environmental sustainability, all of which require increasing agricultural productivity. To increase the productivity, it is crucial to exploit advanced farming (together known as *precision agriculture*) technologies such as yield monitors, remote sensing, and variable rate application, from which site-specific management (SSM) emerges as an effective management strategy [1–4]. This is because SSM can optimize a production system at the subfield level, which amounts to finer-scale optimization than at the field level. Therefore, for both individual profitability and collective societal benefits, SSM has been advocated to farmers over the past two decades [5].

Despite the potential benefits for farmers, the adoption of SSM has been slower than expected [6, 7], which is attributed to a lack of relative advantage over the current management strategies [8], particularly a lack of greater profitability [9, 10]. In principle, those advantages of SSM can be

*saikai@wisc.edu

realized by exploiting advanced technologies (e.g., adjusting fertilizer application rates across the field according to varying soil conditions inferred from satellite imagery). In reality, however, these technologies are quite sophisticated and difficult to exploit. Indeed, many farmers express their concerns about the complexity of these technologies [11] and have yet to be convinced of the value of SSM [12–15]. This lack of actionable procedures or decision support systems for SSM is noted as a serious problem in the literature [16].

The vast majority of SSM research investigates only a handful of management types in particular research environments. Commonly explored types include nitrogen fertilizer [17–24], irrigation water [25, 26], and sowing density [27, 28]. Besides the fact that other types of management (e.g., tillage, spraying, and harvest) are also important, in practice, what really makes farm management difficult is the total number of decisions farmers have to make. Note that the total number of decisions increases exponentially as types of management increases, because each type of management typically involves decisions on its amount, frequency, and timing. Since crop yield and hence the overall effectiveness of SSM are determined by the totality of these management decisions and environmental factors [29–31], independent studies of few management types involving a small number of decisions provide only partial knowledge and can be misleading as complex systems typically involve significant nonlinearity [32, 33].

Besides the small number of management decisions investigated in the existing research, farmers are also concerned about the generalizability of results obtained under the particular research environments, because such results may not be representative of their unique environments [34]. Each farm is unique and operates with different resources (e.g., machinery) in different environments (e.g., soil and weather). Consequently, results from studying "representative" cases are useful for only farmers who face similar managerial and environmental conditions [35, 36]. This problem is exacerbated in modern agriculture where the high dimensionality created by advanced sensor technologies makes it even more difficult for selected cases to be representative, the problem called "the curse of dimensionality" in mathematics [37]. As an example of how particular a typical study setting can be, Lo et al. [24] have studied SSM of nitrogen fertilizer at a university research site in Nebraska that has been "under annual summer corn or soybean production without any tillage and any stover removal" with irrigation water applied through "a center pivot with sprayhead sprinklers positioned every other interrow at a height of 0.6 m above ground" using "GrowSmart Precision Variable Rate Irrigation system". The problem is not how particular existing studies are, as they serve different research purposes. Rather, the problem is a lack of studies that generalize the effects of environmental factors and jointly examine many management variables for the purpose of optimizing SSM in practice.

To address these issues (few management variables only in particular environments) and offer an actionable procedure for SSM, we develop a machine learning algorithm that enables each farmer to efficiently learn unique SSM through on-farm experiments implemented via existing advanced farming technologies. We emphasize the significance of on-farm experimentation [34, 38–40] to deal with both issues. First, on-farm experimentation allows farmers to adaptively design experiments and efficiently navigate a high-dimensional variable space [41]. Second, on-farm experimentation inherently allows each farmer to collect data from his/her unique environment and circumvents the representativeness issue. Particularly for SSM, use of field-scale experimentation is important to deal with spatial, infield variability [36, 42–44]. Notice that our approach is different from the most common type of machine learning (i.e., supervised learning), which constructs an empirical model that assumes specific variables and estimates their associated parameters using a large observational dataset. The constructed model is supposed to be representative of users' production systems and therefore capable of indicating their optimal choices. However, owing to the nature of observational data, such a dataset likely contains insufficient variation in the high-dimensional management space because observed management choices mostly follow the standard recommendations from the existing low-dimensional studies [45]. In other words, purposeful experimentation is crucial to include unconventional management choices and discover unexpected optimal choices. Consequently, it is highly unlikely for such supervised learning models to be able to indicate optimal choices. Our approach is, instead, a machine learning algorithm that allows each farmer to construct a unique model. It is sufficiently versatile to be used for learning optimal SSM through on-farm experiments in a wide range of farming scenarios. The algorithm is based on Bayesian optimization [41], and capable of handling an arbitrary number of management and environmental variables and adapt to unexpected interaction effects. Moreover, if the farmer has historical data, large or small, it can be incorporated as prior knowledge.

In the remainder of the paper, we first mathematically formulate the farmer's problem, and then we describe our machine learning algorithm as a solution to the problem. We test the algorithm's performance and versatility in two simulated environments, with either medium- or high-complexity. To highlight the generality of the approach, we report results on a per-hectare basis so that they can be easily scaled. Results suggest that complex SSM can be learned very efficiently through on-farm experiments within a few years, and it can be 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 precision agriculture equipment for SSM but currently does not implement it due to a lack of knowledge, a typical story about SSM [16]. Suppose that the farmer is now interested in learning optimal SSM 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 existing variable rate technologies, yield monitors, and other sensors that monitor soil properties. Specifically, it is assumed that there is no spatial misalignment among these technologies, all of which has the same spatial resolution. As a result, this resolution defines a site as the observational unit [45], and in each site the farmer collects a pair of data $(\boldsymbol{x}, y)$—applying a vector of management $\boldsymbol{x}$ and observing the corresponding scalar yield y. Let M denote the total number of sites. Site $s \in \{1, 2, \ldots, M\}$ is characterized by a state vector $\boldsymbol{z}_s$, to which management $\boldsymbol{x}_s$ is applied. Then, a site-specific profit function for site s is

$$\pi(\boldsymbol{x}_s|\boldsymbol{z}_s) = py(\boldsymbol{x}_s|\boldsymbol{z}_s) - \boldsymbol{c} \cdot \boldsymbol{x}_s,$$

where $y(\boldsymbol{x}_s|\boldsymbol{z}_s)$ is a site-specific yield function for site s, p is an output price, and $\boldsymbol{c}$ is a vector of prices for $\boldsymbol{x}_s$. Notice that this specification is technically a partial profit, as we subtract only the costs for modeled management $\boldsymbol{x}_s$. Nonetheless, it is immaterial because our analysis is based on the difference between profits using SSM and uniform management.

For conventional low-dimensional yield functions, it is common to use simple concave functions such as quadratic [46–48], negative exponential [49, 50], and piecewise linear [51, 52]. These simple functions may serve well for answering isolated questions about the optimality of a single management variable under homogeneous conditions. 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 site-specific yield function $y(\cdot\,|\,\cdot)$.

Note that each site need not be recognized as distinct or, equivalently, each $\boldsymbol{z}_s$ need not be distinct. A simple consequence of this assumption is that adjacent sites $\{s_1, s_2, \ldots\}$ may have the same values $\boldsymbol{z}_{s_1} = \boldsymbol{z}_{s_2} = \dots$ and form a homogeneous "zone", which receives the same management. The research literature and farmer practice commonly use this type of zone delineation for SSM [26, 53–59]. Our formulation is more general and contains zone delineation as a special case.

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

$$\sum_{s=1}^{M} \pi(\boldsymbol{x}_s|\boldsymbol{z}_s) = \sum_{s=1}^{M} py(\boldsymbol{x}_s|\boldsymbol{z}_s) - \boldsymbol{c} \cdot \boldsymbol{x}_s.$$

The farmer's objective is to learn optimal SSM $\boldsymbol{x}_s^*$ for all $s \in \{1, 2, \dots, M\}$

$$(\boldsymbol{x}_1^*, \dots, \boldsymbol{x}_M^*) = \underset{\boldsymbol{x}_1, \dots, \boldsymbol{x}_M}{\operatorname{argmax}} \sum_{s=1}^{M} \pi(\boldsymbol{x}_s|\boldsymbol{z}_s).$$

In contrast, under uniform management, a single management $\boldsymbol{x}$ is applied to every site s, so that field-level profit function is

$$\sum_{s=1}^{M} \pi(\boldsymbol{x}|\boldsymbol{z}_s) = \sum_{s=1}^{M} py(\boldsymbol{x}|\boldsymbol{z}_s) - \boldsymbol{c} \cdot \boldsymbol{x},$$

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

$$\bar{\boldsymbol{x}}^* = \underset{\boldsymbol{x}}{\operatorname{argmax}} \sum_{s=1}^{M} \pi(\boldsymbol{x}|\boldsymbol{z}_s).$$

2.2 Solution algorithm

We construct an algorithm based on Bayesian optimization (BO) [60, 61], a class of numerical optimization techniques to find the global optimum of an unknown objective function. As with many other numerical optimization techniques, BO navigates the search space by examining one 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. A Gaussian process (GP) statistical model is the standard choice in the literature.

BO has two features that make it suitable for agricultural experiments [41]. First, GP as a nonparametric Bayesian model is sufficiently flexible so that it can adapt to cases in which the objective function takes a complex shape. In high-dimensional SSM, this complexity will likely happen due to strong interactions among the 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, sample efficiency is a strongly desirable feature.

While the basic BO sequentially processes one sample at a time, we modify it using the "batch BO" proposed by Saikai et al. [41] in order to process M samples at a time. Each year, the algorithm proceeds as follows:

1. Prescribe $\boldsymbol{x}_s$ for each site s by maximizing the acquisition function $\alpha(\boldsymbol{x}|\boldsymbol{z}_s)$
2. Observe a yield $y(\boldsymbol{x}_s|\boldsymbol{z}_s)$ for each s
3. Compute the corresponding $\pi(\boldsymbol{x}_s|\boldsymbol{z}_s)$ for each s
4. Update the GP with $\{(\boldsymbol{x}_s, \boldsymbol{z}_s, \pi_s)\}_{s=1}^{M}$ and the samples from the preceding years.

After completing the planned number of years of experiments, a candidate for $\boldsymbol{x}_s^*$ for each s can be obtained by maximizing the mean function of the learned GP with fixed $\boldsymbol{z}_s$. Below is the complete algorithm.

Algorithm 1 Batch Bayesian optimization for site-specific management

1: **require:** T, M, $\mathcal{S}$, GP, α
2: **for** $t \in \{1, 2, \ldots, T\}$ **do**
3:     $\mathcal{I} \leftarrow \{\ \}$
4:     $\widehat{GP} \leftarrow GP$
5:     **for** $s \in \{1, 2, \ldots, M\}$ **do**
6:         $\boldsymbol{x}_s \leftarrow \mathrm{argmax}_{\boldsymbol{x}}\, \alpha(\boldsymbol{x}|\boldsymbol{z}_s)$
7:         $\mathcal{I} \leftarrow \mathcal{I} \cup \{(\boldsymbol{x}_s, \boldsymbol{z}_s, \underline{\pi})\}$ where $\underline{\pi} = \min\{\mathcal{S}_\pi\}$
8:         Update $\widehat{GP}$ with $\mathcal{S} \cup \mathcal{I}$
9:     **for** $s \in \{1, 2, \ldots, M\}$ **do**
10:         $y_s \leftarrow Oracle(\boldsymbol{x}_s|\boldsymbol{z}_s)$
11:         $\pi_s \leftarrow py_s - \boldsymbol{c} \cdot \boldsymbol{x}_s$
12:         $\mathcal{S} \leftarrow \mathcal{S} \cup \{(\boldsymbol{x}_s, \boldsymbol{z}_s, \pi_s)\}$
13:     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, $\alpha(\,\cdot\,|\boldsymbol{z}_s)$ is an acquisition function for site s defined based on the Gaussian process GP, and $Oracle(\boldsymbol{x}|\boldsymbol{z})$ returns an observed yield when $\boldsymbol{x}$ is applied to a site characterized by $\boldsymbol{z}$. As in Saikai et al. [41], for the acquisition function, we use the expected improvement [62, 63]. Also, for the oracle function, we use a crop simulator as described in the next section. 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 sites are numbered. Another detail is that, when updating GP, we fit the hyperparameters of the 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 [64] widely used for various purposes, including generating synthetic datasets [65–71]. In each environment, we run the algorithm to learn optimal SSM over T years and compare the profit resulting from implementing the learned SSM against the benchmark profit resulting from uniform management. This analysis assumes that uniform management follows university extension recommendations.

Depending on specific scenarios, we can customize the algorithm in many ways. An interesting modification is to incorporate observational data collected prior to beginning on-farm experimentation, the case for many farmers. When using uniform management as a benchmark for comparison, a natural dataset incorporated is the data from implementing uniform management, as it represents the existing knowledge. Here, we initialize the GP embedded in the algorithm as follows: let $\{(\bar{\boldsymbol{x}}, \boldsymbol{z}_s, \bar{\pi}_s)\}_{s=1}^{M}$ be a set of the uniform management ($\bar{\boldsymbol{x}}$), site characteristics ($\boldsymbol{z}_s$), and the corresponding profits ($\bar{\pi}_s$) and then, we fit the GP to these M data points before the algorithm starts a learning process. Note that the use of uniform management for both benchmark and prior knowledge is a useful case for 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 100 Monte Carlo samples.

Though the algorithm's applicability is by no means restricted to the scenarios described in this section, we rely on the APSIM simulator and construct illustrative test beds within its capability. We simulate a maize production system in Ames, Iowa using the daily weather data for 2013, the most recent annual dataset available in APSIM. In terms of management variables, we follow Saikai et al. [41] and identify six variables $\boldsymbol{x} = (x^1, \ldots, x^6)$ in the APSIM maize module:

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

Based on Iowa State University extension recommendations, we specify uniform management ($\bar{\boldsymbol{x}}$) as:

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

$\bar{x}^1 = 8$, $\bar{x}^2 = 50$, and $\bar{x}^3 = 0.76$ follow from Farnham [72] and Elmore [73]. The recommended total nitrogen amount is identified by using the Corn Nitrogen Rate Calculator [74], which gives $\bar{x}^4 + \bar{x}^5 + \bar{x}^6 = 201$. We evenly split it into $\bar{x}^4 = \bar{x}^5 = \bar{x}^6 = 67$. Finally, for calculating profits, the output price is $p = \$0.177/\text{kg}$ [75], and input costs are $c^1 = \$3.64/1000$ seeds and $c^4 = c^5 = c^6 = \$1.29/\text{kg}$ [76]. We assume no cost for sowing depth and row spacing, which implies the cost vector $\boldsymbol{c} = (c^1, 0, 0, c^4, c^5, c^6)$.

Based on these same sources, when the algorithm searches for the optimal management, the search space is restricted 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, data points resulting from the benchmark uniform management $\{(\bar{\boldsymbol{x}}, \boldsymbol{z}_s, \bar{\pi}_s)\}_{s=1}^{M}$ are the only existing dataset incorporated prior to beginning on-farm experimentation. Since initial uniform management provides no variation in $\boldsymbol{x}$, to build up smoothly, in the first year, the algorithm randomly chooses $\boldsymbol{x}$ for each s from the range defined by $\pm 50\%$ of the uniform management.

2.3.1 Scenario A (medium complexity)

This scenario assumes that a square field is divided into a grid of 25 sites ($M = 25$). All the sites are distinct, each characterized by a state vector $\boldsymbol{z}_s = (z_s^1, z_s^2)$ 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 $z_s^2 \in \{2.56, 2.88, 3.2, 3.52, 3.84\}$ (0%, $\pm 10\%$ and $\pm 20\%$ from the default values in the APSIM soil module used).

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 1: Simulated maize field divided into a grid of 25 distinct sites. The first number in each grid indicates plant available water capacity (mm) and the second indicates organic carbon (%) at that site.

2.3.2 Scenario B (high complexity)

This scenario imagines that the farmer possesses more precise equipment that can operate at a more granular scale, and so divides a field into more granular sites: $16 \times 9 = 144$ sites. We also assume that the farmer has conducted more exhaustive soil tests, measuring four state variables (z^1, z^2, z^3, z^4) in each site:

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

The addition of z^3 and z^4 is because of their significance for nitrogen management [77] 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 [78], so that SSM in scenario B is more complex than in scenario A. We generate a state vector for each site in a random but spatially correlated fashion, namely, a random walk (see Appendices for details). Below are summary statistics of the generated state vectors for the 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 reporting four numbers at each site, we illustrate the in-field variability using a yield map arising from applying the uniform management to the generated field. Since each site receives the same management, the variability in yield indicates the variability in the underlying growing conditions.

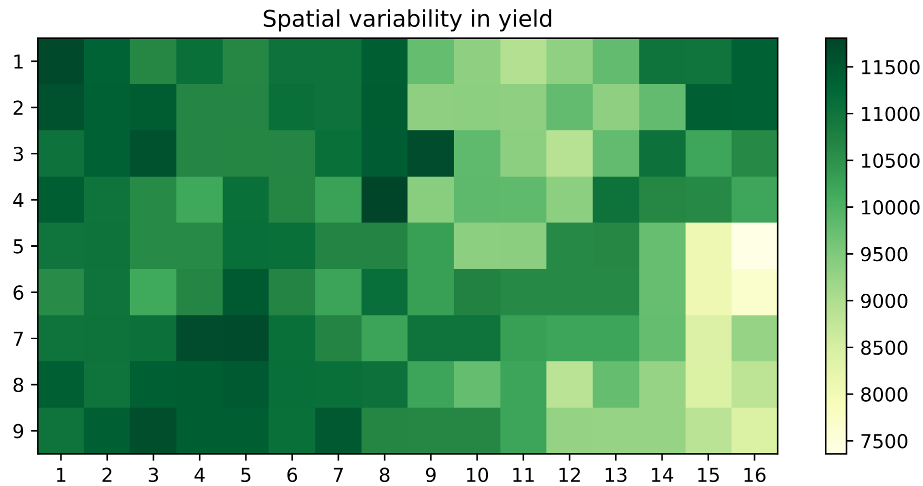


Figure 2: Yield map (kg/ha) resulting from the uniform management in scenario B. Axis ticks indicate site coordinates.

3 Results

3.1 Scenario A (medium complexity)

Table 1 and Figure 3 report field-level profits (\$/ha) from implementing the SSM learned after conducting experiments for T years. Specifically, the value for each $T \in \{1, 2, \ldots, 10\}$ is the profit if the farmer terminates the experiments after T years and implements the learned SSM without further improvement.

Years (T)	1	2	3	4	5	6	7	8	9	10
Learned	1103	1237	1266	1274	1277	1280	1283	1284	1285	1285
Uniform	1234	1234	1234	1234	1234	1234	1234	1234	1234	1234
Difference	-131	3	32	40	43	46	49	50	51	51
										(\$/ha)

Table 1: Field-level profits (\$/ha) from implementing the learned SSM and uniform management in scenario A.

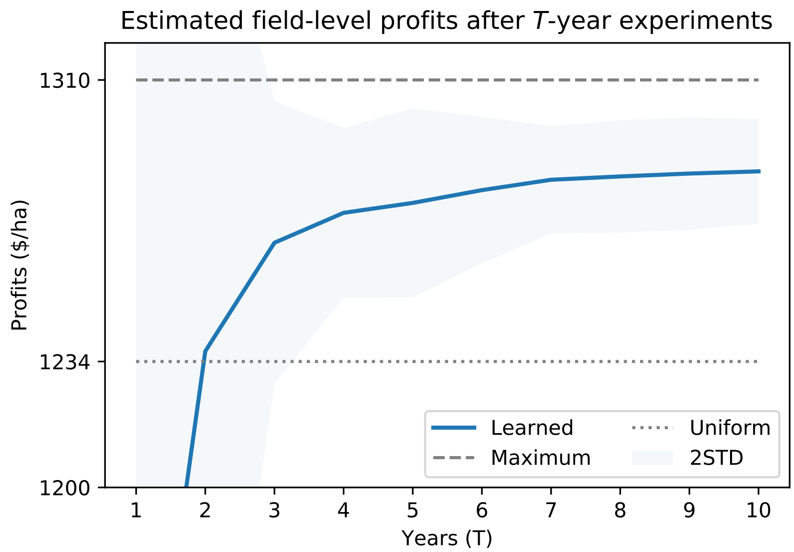


Figure 3: Learning curve of the algorithm with profits plotted against years of experiments in scenario A. The shaded areas indicate two standard deviations around each mean field-level profit over 100 Monte Carlo samples.

The shaded areas indicate two standard deviations around each mean field-level profit over 100 Monte Carlo samples. The dashed line indicates the maximum possible profit obtained by implementing the optimal SSM, while the dotted line indicates the profit from uniform management.

Since a field-level profit is the sum of the site-specific profits, we next examine profit at each site. Figure 4 illustrates profits from the learned SSM after five years and uniform management. $T = 5$ is chosen because the learning mostly levels off and the deviation from the mean prediction becomes small after four or five years.

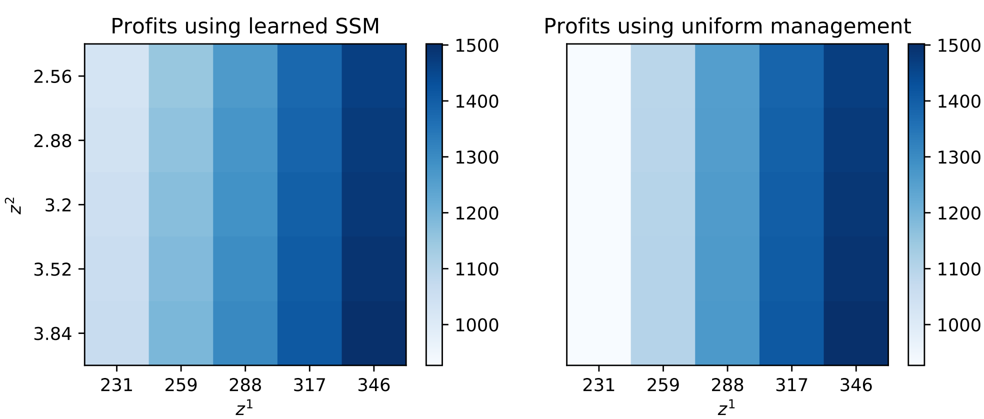


Figure 4: Site-specific profits ($/ha) for the learned SSM after 5 years and uniform management in scenario A

As indicated in the panel for uniform management, plant available water capacity (z^1) has a much stronger influence on profits than organic carbon (z^2). However, as both variables increase, the site becomes more fertile (though hard to see for organic carbon z^2). Since panels for both management systems look quite similar, to highlight their difference, Figure 5 illustrates the difference at each site.

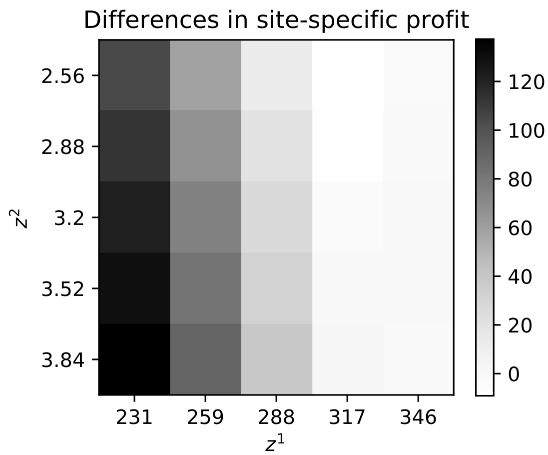


Figure 5: Differences in site-specific profit (\$/ha) for SSM versus uniform management in scenario A. The maximum difference is \$138/ha at site (231,3.84) and the minimum difference is \$-9.4/ha at site (317,2.56).

Finally, Figure 6 reports the learned SSM $(x^1, \ldots, x^6)$ after 5 years.

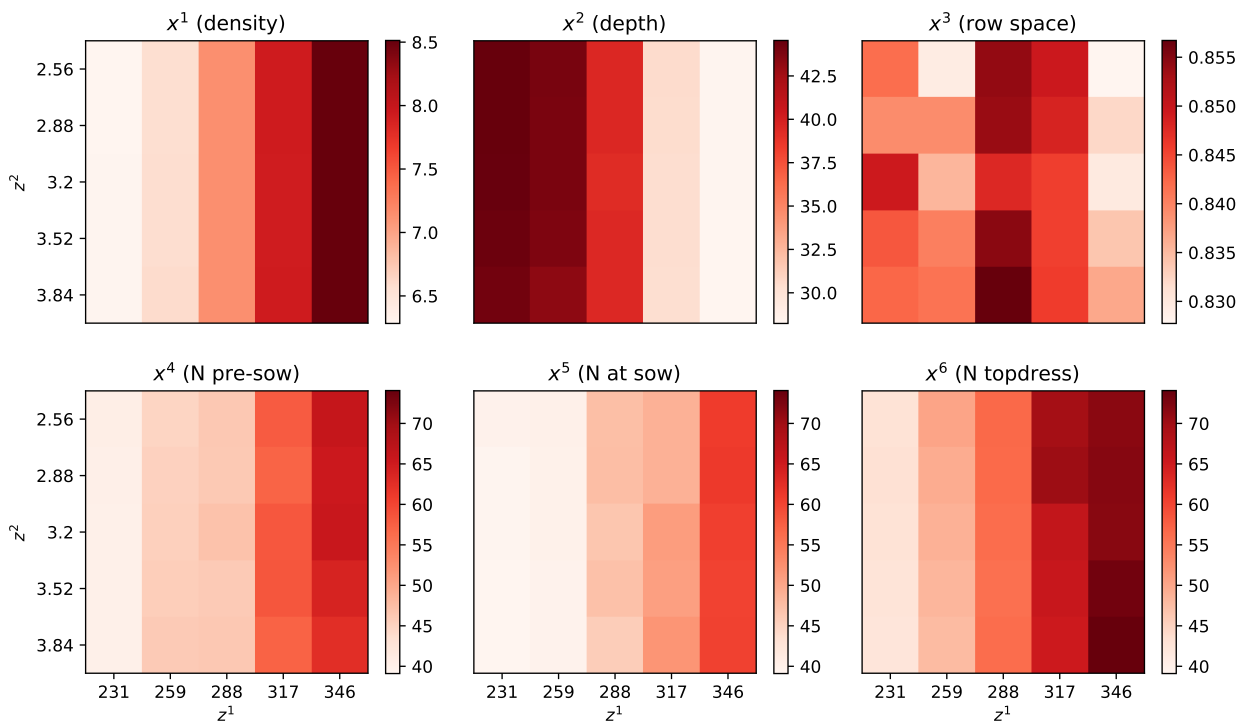


Figure 6: Learned SSM after 5 years in scenario A. The average sowing density is 7.3 seeds/m^2, and the average amount of total nitrogen is 156 kg/ha.

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 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)

Table 2 and Figure 7 report field-level profits (\$/ha) from implementing the SSM learned after conducting experiments for T years. Again, the value for each $T \in \{1, 2, \ldots, 10\}$ is the profit if the farmer terminates the experiments after T years and implements the learned SSM without further improvement.

Years (T)	1	2	3	4	5	6	7	8	9	10
Learned	1319	1324	1331	1333	1335	1335	1336	1337	1338	1339
Uniform	1295	1295	1295	1295	1295	1295	1295	1295	1295	1295
Difference	24	29	36	38	40	40	41	42	43	44
										(\$/ha)

Table 2: Field-level profits (\$/ha) from implementing the learned SSM and uniform management in scenario B.

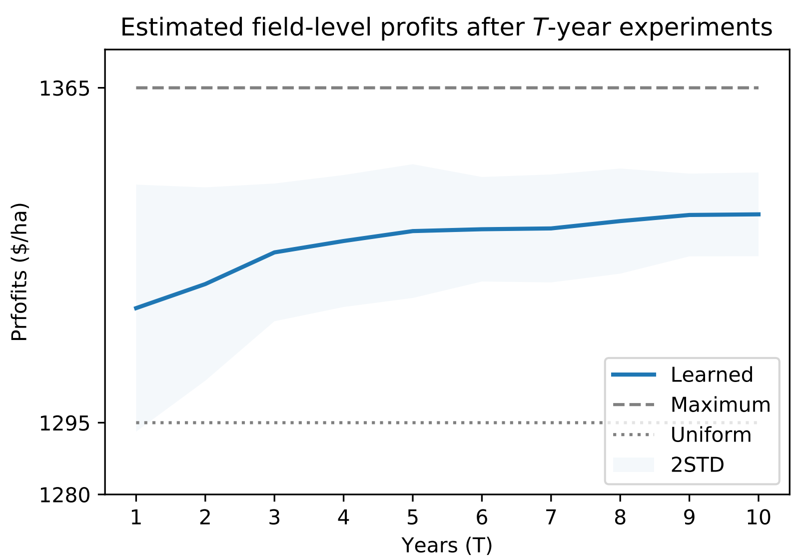


Figure 7: Learning curve of the algorithm with profits plotted against years of experiments in scenario B. The shaded areas indicate two standard deviations around each mean field-level profit over 100 Monte Carlo samples.

The shaded areas indicate two standard deviations around each mean field-level profit over 100 Monte Carlo samples. Again, the dashed line indicates the maximum possible profit obtained by implementing the optimal SSM, while the dotted line indicates the profit from uniform management.

The following heatmaps (Figure 8) compare the site-specific profits from the learned SSM at $T = 5$ and uniform management.

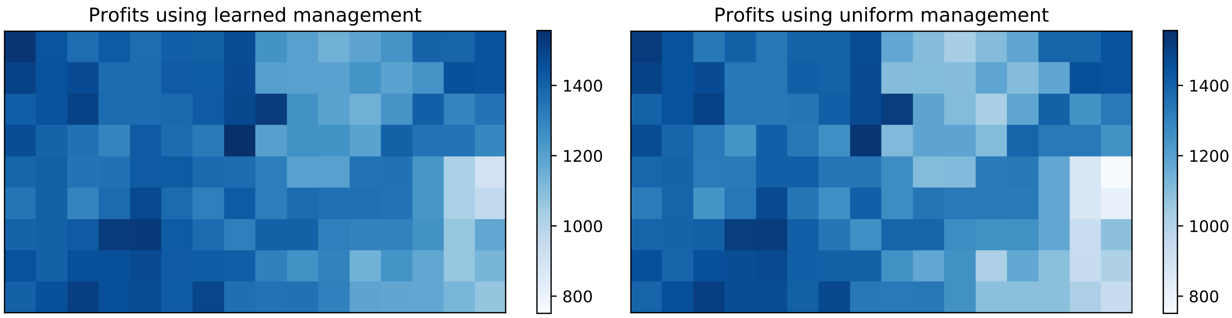


Figure 8: Site-specific profits (\$/ha) for the learned SSM after 5 years and uniform management in scenario B

As in scenario A, the difference between the two management systems is difficult to discern. The SSM, however, has higher profits (darker colors) in low-yielding sites. Figure 9 illustrates the difference at each site.

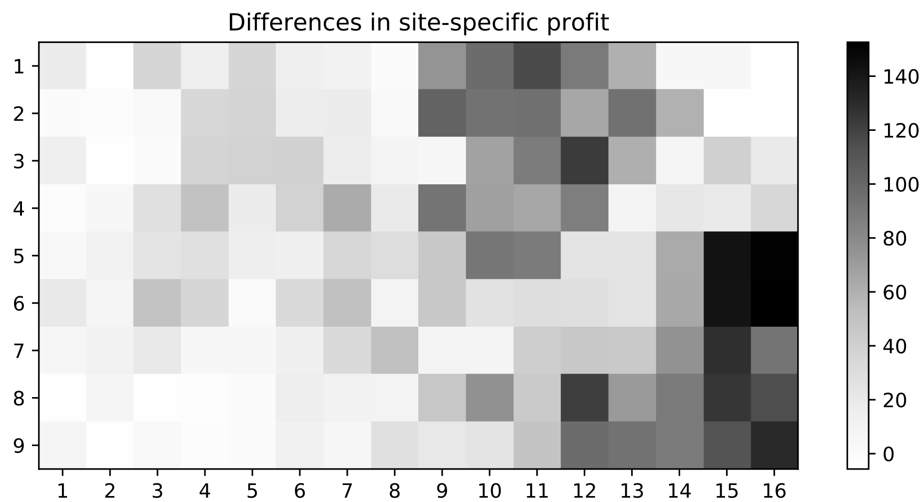
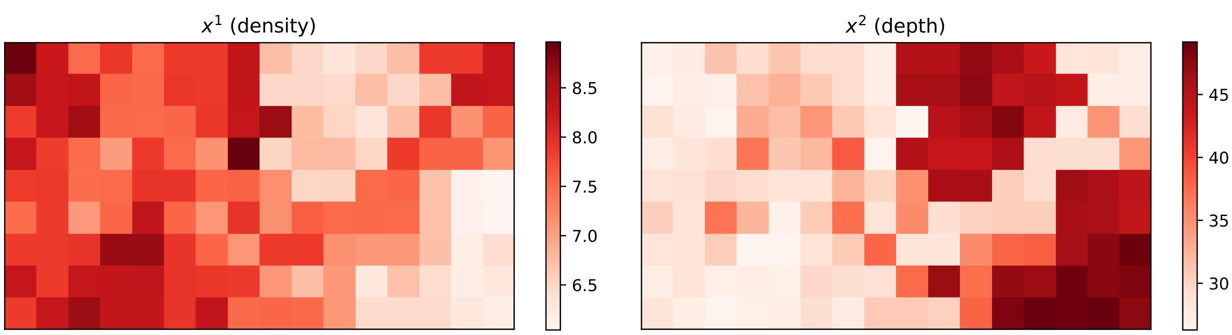


Figure 9: Differences in site-specific profit (\$/ha) for SSM versus uniform management in scenario B. The maximum difference is \$153/ha at site (16,6) and the minimum difference is \$-5.7/ha at site (8,1).

Finally, Figure 10 illustrates the learned SSM $(x^1, \ldots, x^6)$ after 5 years.



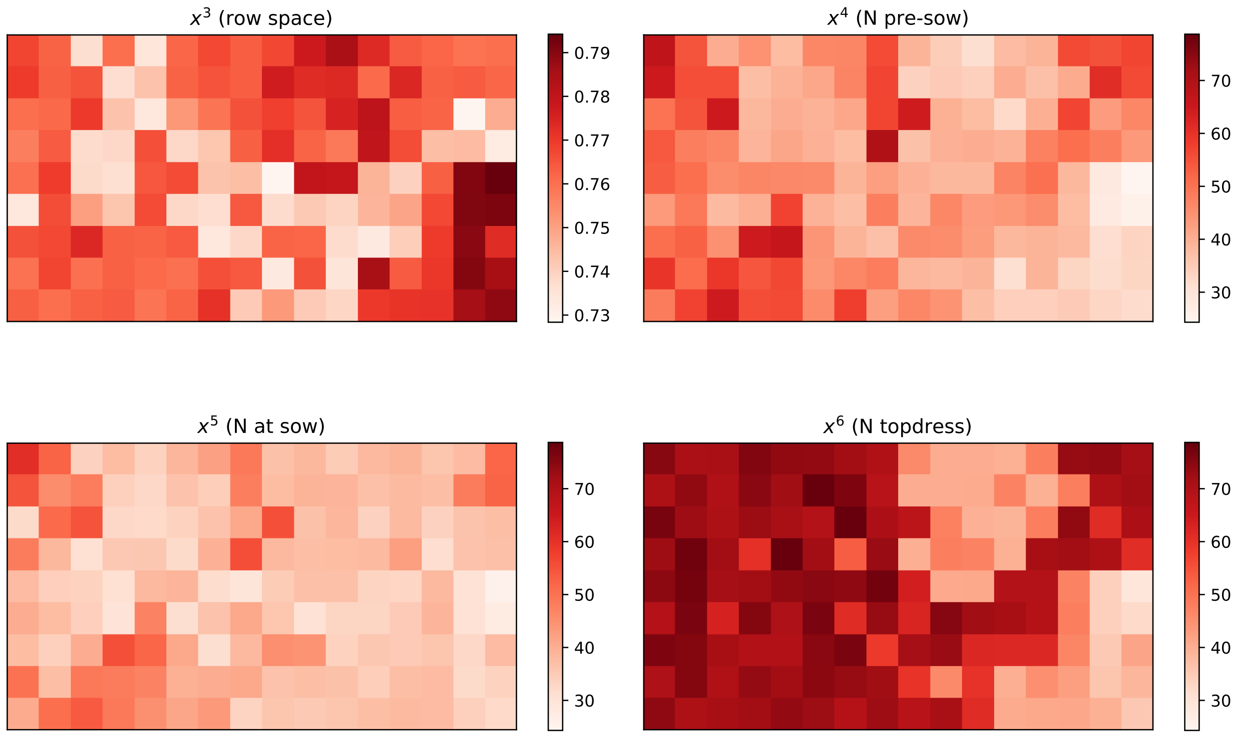


Figure 10: Learned SSM after 5 years in scenario B. The average sowing density is 7.4 seeds/m^2, and the average amount of total nitrogen is 146 kg/ha.

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

$$\frac{1}{144}\sum_{s=1}^{144}\sum_{i=4}^{6} x_s^i = 146 \text{ kg/ha}.$$

As a result, \$40/ha higher profit is achieved by using 0.6 fewer seeds/m^2 and 55 kg/ha less nitrogen than uniform management.

4 Discussion

The results for field-level profit presented in Tables 1 and 2 and Figures 3 and 7 are well aligned with our common notion about learning—the longer an algorithm learns, the higher profit the learned SSM generates. When the algorithm starts with little existing data to incorporate, the algorithm has difficulty identifying good management. This is particularly the case in scenario A, in which the profit from the learned SSM after one year is far below that for uniform management. In contrast to scenario A, good management is found in scenario B even after one year because, although most sites are technically distinct, some sites are similar to each other and provide mutual information. As a result, with a greater number of sites in the field, more information is collected each year. Despite the low performance in the first year in scenario A, the algorithm quickly learns and its performance surpasses the performance with uniform management after two years. In both scenarios, the algorithm continues to learn and widen the performance gap. With five years of learning, the estimated profit reaches \$1,277 or 97% of the maximum possible profit (\$1,310) in scenario A and \$1,335 or 98% of the maximum possible profit (\$1,365) in scenario B.

The random sampling used in year 1 creates the large shaded area formed by two standard deviations around the mean predictions over the first few years in scenario A (Figure 3), indicating that imprecise prediction of mean profits. However, performance dramatically improves after four years, and thereafter its spread around the mean profits 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, the algorithm consistently learns good SSM. Scenario B exhibits far less imprecision due to the larger sample size used right from the beginning (Figure 7).

As seen in Figures 4,5,8, and 9, the higher field-level profits for the learned SSM are due mainly to their higher profits from the low-yielding sites (e.g., sites with $z^1 \in \{231, 259\}$ in scenario A and sites around (11,2) and (15,7) in scenario B). These results imply that uniform management is excessively tailored to the high-yielding conditions—$z^1 \in \{317, 346\}$ in scenario A and the left half of the field in scenario B—leading to the decrease in profitability in the low-yielding sites where it is optimal to put less inputs. Overall, albeit not necessarily true in other environments, the algorithm discovers that it is profitable to put more inputs in the high-yielding sites and less in the low-yielding sites as indicated in panels for x^1, x^4, x^5, and x^6 in Figure 6 and 10.

We dismiss the patchy look of row spacing (x^3) in Figures 6 and 10 as an artifact of numerical optimization, which strictly distinguishes two values whenever one results in even a minuscule amount greater than the other. Indeed, the color bar for x^3 has a very small range (0.830–0.855 in scenario A and 0.73–0.79 in scenario B), indicating little practical significance for learning SSM. Nonetheless, we have included row spacing in the learning of SSM because we do not assume its insignificance before running the algorithm. In general, with little prior knowledge about which management variables are significant and should be included for learning their optimal management choices, we should include them and let the algorithm learn.

In both scenarios, the learned SSM is evidently more efficient in input use for generating profits than the benchmark uniform management. After five years, for example, the learned SSM generates \$43/ha higher profits with 45 kg/ha less nitrogen in scenario A and \$40/ha higher profits with 55 kg/ha less nitrogen in scenario B. In terms of yield, the learned management produces 210 kg/ha less maize in scenario A and 278 kg/ha less maize in scenario B. While the SSM optimization is guided by profit maximization, it turns out to be environmentally more sustainable as well because both costs of fertilizer (i.e., to profitability and to the environment) are aligned so that less is better. However, higher yield does not necessarily coincide with higher profit, though yield increases with more inputs, substantially higher input costs can reduce profit.

Although the focus of this paper is on development and demonstration of the algorithm, we mention some implications and implementation of the algorithm in practice for future empirical studies. Imagine a corn farmer in Iowa whose farming situation is well captured by scenario B. Then, implications of the higher projected profits largely depend on whether the farmer needs to invest in new equipment and if so, how large the operation scale is. As mentioned in section 2.1, a typical story about SSM is that, despite the existing access to required equipment, a lack of actionable procedure for SSM prevents the farmer from implementing it [16]. In this case, extra profits may be sufficient to cover costs for experiment and overcome psychological barriers to change (e.g. accepting lower yield for higher profit). In contrast, if new investment is necessary, further financial analysis is required. Suppose that the farm size is 100 ha and extra \$40/ha is projected. Then, the analysis involves comparing a stream of extra \$4,000/year against equipment and other financial costs. Finally, the algorithm can be implemented in practice by writing its recommended management choices into a suitable prescription file whose format is specified by the farmer's equipment and software. For example, we may write into a format required by Ag Leader SMS$^{\text{TM}}$, popular software used in many farmers in the U.S. to communicate with a wide range of precision farming equipment.

Despite the promising results, there are several clarifications and limitations to note before real-world implementations, as well as future research needs.

- While we have formulated a farmer's problem as profit maximization and developed an algorithm to solve it, in reality, many farmers are concerned with maximizing their yields. In fact, our optimization framework is flexible and can be applied to solving yield maximization problems as well, which will be the focus in future work. In this paper, we concentrate on the profit maximization formulation for the following reasons. First and foremost, since the main objective of this paper is to introduce the novel optimization method to the literature, we try to highlight its features and usefulness for precision agriculture. To this end, profit maximization formulation is technically simpler than yield maximization. In addition, as extensively studied over the past two decades, the adoption of PA (or the lack thereof) is explained by many and complex determinants, among which profitability is identified as one of the key factors in the literature [7, 8, 10, 79]. Even for those who try to maximize yields, as part of commercial enterprises, yield maximization is rarely unconstrained. Indeed, there are typically implicit upper bound for input use (e.g., avoiding unnecessary fertilizer application). In this case, constrained yield maximization becomes similar to profit maximization.
- This analysis assumes no costs for switching management from site to site, which can be unrealistic for some inputs and management types. For example, varying types of fertilizer, seed treatments or hybrids may require equipment modifications, multiple field passes or additional labor. As another example, changing seeding rates too frequently may put an excessive strain on and damage an electric motor. If these are costs to consider, optimization of SSM will be even harder due to the switching frequencies of management as another set of control variables to optimize.
- The algorithm assumes choice variables are continuous, even though choices of continuous variables can be constrained for various practical reasons. For example, though the algorithm may recommend fertilizer application rates that differ by less than 0.1 kg/ha, such small differences are impossible to implement practically with current equipment. Real-world implementation will require modifying the algorithm to convert continuous choice variables into appropriate discrete variables.
- The simulations use a single season's weather pattern, output price, and input prices that prevail over T years. In reality, these differ from year to year, and such fluctuations may have strong implications for the algorithm's performance.
- The current algorithm assumes risk neutrality of farmers when the acquisition function prescribes the next sampling choices. In the real world, however, many farmers are risk averse [80], for whom it will be difficult to accept and implement some strongly explorative prescriptions made by the algorithm that reflects expected profit but not its variance. To be more realistic and useful for practical applications, future work needs to modify the acquisition function and parameterize a level of risk aversion so that the algorithm can be uniquely adjusted to each farmer.
- This approach is limited to one-shot optimization in which the farmer makes all the management decisions at the beginning of the year, implements them, and waits to see results at the end of the year. In practice, many management choices are sequentially made throughout the year, while they affect site characteristics and are affected by evolving site characteristics.

To handle these more realistic situations requires dynamic models with information feedback and learning that take place both within and across years.

- The simulations assume no spatial interaction between input choices across sites (e.g., fertilizer use on one site does not affect adjacent sites). Depending on the grid granularity, these assumptions may be excessively strong under some circumstances. Thus, besides the temporal dynamics mentioned above, future work focuses on developing spatiotemporally explicit models that handle dynamics over space and time.
- As with many other machine learning studies on large and complex datasets, it is extremely beneficial to have access to a simulator of high fidelity. We find in APSIM only six management variables and four environmental variables suitable for this study. Since the algorithm is very flexible and capable of dealing with an arbitrary number of variables, demonstration of the algorithm would be more realistic and compelling with a simulator in which a greater number of variables and resulting yield are intricately interdependent, better representing the complexity of biophysical systems. Thus, further advancement of crop simulators is crucial for validation and improvement of the current algorithm.

5 Conclusions

We have proposed an algorithmic approach to optimizing complex site-specific management with many management and environmental variables. The proposed algorithm enables individual farmers to efficiently learn their own site-specific management through on-farm experiments. We have demonstrated its performance using simulated environments. The results have provided a positive answer to both the learnability of complex site-specific management and the higher profitability possible relative to uniform management. While these results are promising, we do not know in general under which environment the algorithm works. Thus, we need more validation studies, especially field experiments with farmer cooperators. Long-term, the results suggest that on-farm experimentation implemented with precision agriculture equipment can help to realize the benefits of precision agriculture—more profitable management, greater food security, and improved environmental sustainability.

References

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

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

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

[4] Hermann Auernhammer. "Precision farming—the environmental challenge". In: *Computers and electronics in agriculture* 30.1 (2001), pp. 31–43.

[5] Simon Cook and R. G. V. Bramley. "Precision agriculture—opportunities, benefits and pitfalls of site-specific crop management in Australia". In: *Australian Journal of Experimental Agriculture* 38.7 (1998), pp. 753–763.

[6] R. G. V. Bramley. “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 (2009), pp. 197–217.

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

[8] Hari Sharan Pathak, Philip Brown, and Talitha Best. “A systematic literature review of the factors affecting the precision agriculture adoption process”. In: *Precision Agriculture* (2019).

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

[10] Markus Gandorfer and Andreas Meyer-Aurich. “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. Springer International Publishing, 2017, pp. 79–92.

[11] R. G. V. Bramley and J. Ouzman. “Farmer attitudes to the use of sensors and automation in fertilizer decision-making: nitrogen fertilization in the Australian grains sector”. In: *Precision Agriculture* 20.1 (2019), pp. 157–175.

[12] Emma Leonard, Rohan Rainbow, A Laurie, David Lamb, R Llewellyn, Ed Perrett, Jay Sanderson, Andrew Skinner, T Stollery, and Leanne Wiseman. *Accelerating precision agriculture to decision agriculture: Enabling digital agriculture in Australia*. Australi: Cotton Research and Development Corporation, 2017.

[13] John M Antle, James W Jones, and Cynthia E Rosenzweig. “Next generation agricultural system data, models and knowledge products: Introduction”. In: *Agricultural Systems* 155 (2017), pp. 186–190.

[14] John M Antle. “Data, Economics and Computational Agricultural Science”. In: *American Journal of Agricultural Economics* 101.2 (2019), pp. 365–382.

[15] Jess Lowenberg-DeBoer and Bruce Erickson. “Setting the record straight on precision agriculture adoption”. In: *Agronomy Journal* (2019).

[16] Jessica Lindblom, Christina Lundström, Magnus Ljung, and Anders Jonsson. “Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies”. In: *Precision Agriculture* 18.3 (2017), pp. 309–331.

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

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

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

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

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

[22] Ingo Pahlmann, Ulf Böttcher, and Henning Kage. “Developing and testing an algorithm for site-specific N fertilization of winter oilseed rape”. In: *Computers and Electronics in Agriculture* 136 (2017), pp. 228–237.

[23] Dave Pannell, Markus Gandorfer, and Alfons Weersink. “How flat is flat? Measuring payoff functions and the implications for site-specific crop management”. In: *Computers and Electronics in Agriculture* 162 (2019), pp. 459–465.

[24] Tsz Him Lo, Daran R. Rudnick, Brian T. Krienke, Derek M. Heeren, Yufeng Ge, and Tim M. Shaver. “Water effects on optical canopy sensing for late-season site-specific nitrogen management of maize”. In: *Computers and Electronics in Agriculture* 162 (2019), pp. 154–164.

[25] Amir Haghverdi, Brian G. Leib, Robert A. Washington-Allen, Michael J. Buschermohle, and Paul D. Ayers. “Studying uniform and variable rate center pivot irrigation strategies with the aid of site-specific water production functions”. In: *Computers and Electronics in Agriculture* 123 (2016), pp. 327–340.

[26] Nestor M. Cid-Garcia, Angel G. Bravo-Lozano, and Yasmin A. Rios-Solis. “A crop planning and real-time irrigation method based on site-specific management zones and linear programming”. In: *Computers and Electronics in Agriculture* 107 (2014), pp. 20–28.

[27] Xiantao He, Youqiang Ding, Dongxing Zhang, Li Yang, Tao Cui, and Xiangjun Zhong. “Development of a variable-rate seeding control system for corn planters Part II: Field performance”. In: *Computers and Electronics in Agriculture* 162 (2019), pp. 309–317.

[28] Xiantao He, Youqiang Ding, Dongxing Zhang, Li Yang, Tao Cui, and Xiangjun Zhong. “Development of a variable-rate seeding control system for corn planters Part I: Design and laboratory experiment”. In: *Computers and Electronics in Agriculture* 162 (2019), pp. 318–327.

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

[30] David S Bullock, Jess Lowenberg-DeBoer, and Scott M Swinton. “Adding value to spatially managed inputs by understanding site-specific yield response”. In: *Agricultural Economics* 27.3 (2002), pp. 233–245.

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

[32] Miguel A Altieri. *Agroecology: The science of sustainable agriculture*. 2nd. CRC Press, 2018.

[33] Stephen R. Gliessman. “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, 1990, pp. 3–10.

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

[35] Kwanchai A Gomez and Arturo A Gomez. *Statistical procedures for agricultural research*. John Wiley & Sons, 1984.

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

[37] Richard E Bellman. *Adaptive control processes: a guided tour*. Vol. 2045. Princeton university press, 2015.

[38] Terry Griffin. *Collating and analysing small data to make big decisions – Can it improve farm productivity and profitability?* Grain Research and Development Corporation. 2018. 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.

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

[40] Andreas Meyer-Aurich, Terry W. Griffin, Ruprecht Herbst, Antje Giebel, and Nawaz Muhammad. "Spatial econometric analysis of a field-scale site-specific nitrogen fertilizer experiment on wheat (Triticum aestuvum L.) yield and quality". In: *Computers and Electronics in Agriculture* 74.1 (2010), pp. 73–79.

[41] Yuji Saikai, Vivak Patel, Shawn P Conley, and Paul D Mitchell. *Adaptive experimental design using Bayesian optimization to improve the cost efficiency of small-plot field trials*. 2020. URL: https://github.com/ysaikai/AEDBO.

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

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

[44] R. G. V. Bramley, KJ Evans, KJ Dunne, and DL Gobbett. "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 (2011), pp. 341–350.

[45] David S Bullock, Maria Boerngen, Haiying Tao, Bruce Maxwell, Joe D Luck, Luciano Shiratsuchi, Laila Puntel, and Nicolas F Martin. "The data-intensive farm management project: Changing agronomic research through On-farm precision experimentation". In: *Agronomy Journal* (2019).

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

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

[48] B M Whelan, J A Taylor, and A B McBratney. "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 (2012), pp. 47–56.

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

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

[51] Frederic Ouedraogo and B Wade Brorsen. "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 (2018), pp. 87–102.

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

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

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

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

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

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

[58] Adriana Gili, Cristian Álvarez, Ramiro Bagnato, and Elke Noellemeyer. "Comparison of three methods for delineating management zones for site-specific crop management". In: *Computers and Electronics in Agriculture* 139 (2017), pp. 213–223.

[59] Elia Scudiero, Pietro Teatini, Dennis L. Corwin, Rita Deiana, Antonio Berti, and Francesco Morari. "Delineation of site-specific management units in a saline region at the Venice Lagoon margin, Italy, using soil reflectance and apparent electrical conductivity". In: *Computers and Electronics in Agriculture* 99 (2013), pp. 54–64.

[60] Eric Brochu, Vlad M Cora, and Nando De Freitas. "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* (2010).

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

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

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

[64] Dean P Holzworth, 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. "APSIM–evolution towards a new generation of agricultural systems simulation". In: *Environmental Modelling & Software* 62 (2014), pp. 327–350.

[65] Zhenong Jin, Sotirios V Archontoulis, and David B Lobell. "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 (2019), pp. 12–22.

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

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

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

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

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

[71] Marshall Burke and David B Lobell. "Satellite-based assessment of yield variation and its determinants in smallholder African systems". In: *Proceedings of the National Academy of Sciences* 114.9 (2017), pp. 2189–2194.

[72] Dale Farnham. *Corn Planting Guide*. 2001.

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

[74] John Sawyer. *The Corn Nitrogen Rate Calculator*. 2019. URL: http://cnrc.agron.iastate.edu/nRate.aspx.

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

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

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

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

[79] Michael H Castle, Bradley D Lubben, Joe D Luck, and Taro Mieno. *Precision Agriculture Adoption and Profitability*. Agricultural Economics, University of Nebraska–Lincoln. 2017. URL: https://agecon.unl.edu/cornhusker-economics/2017/precision-agriculture-adoption-profitability.

[80] M Monjardino, T McBeath, J Ouzman, Rick Llewellyn, and B Jones. "Farmer risk-aversion limits closure of yield and profit gaps: A study of nitrogen management in the southern Australian wheatbelt". In: *Agricultural Systems* 137 (2015). Publisher: Elsevier, pp. 108–118.

[81] Carl Edward Rasmussen and Christopher K Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

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

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

Appendices

APSIM configuration

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

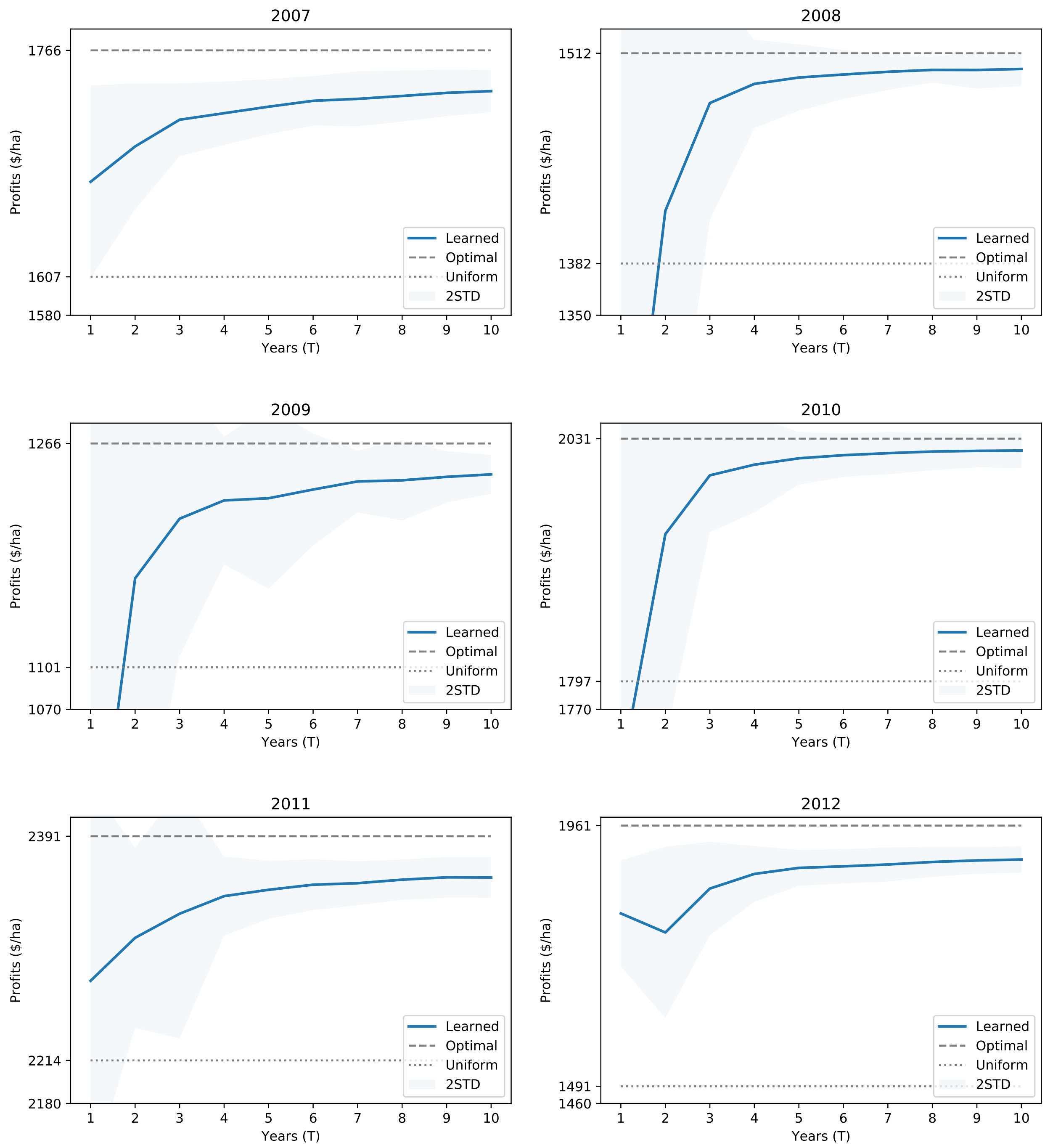
- Metfile: USA_Iowa_Ames.met
- Calendar: Jan 1, 2013 - Dec 31, 2013
- Cultivar: Pioneer 3394
- Sowing window START data: 15-apr
- Sowing window END data: 2-may
- Soil: Iowa Nicollet soil series
- Initial nitrogen: 0 kg/ha for both NO3 and NH4 for scenario A
- Initial water: 80% filled from top

Sensitivity analysis

In addition to two simulated environments with medium- 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 [75, 76].

Year	2007	2008	2009	2010	2011	2012
Output price ($/kg)	0.17	0.16	0.14	0.21	0.24	0.27
Seed price ($/1000 seeds)	1.82	2.10	3.13	3.44	3.25	3.40
Nitrogen price ($/kg)	0.69	1.02	1.51	0.73	1.13	1.40

Note that all sensitivity analysis was conducted under the environments with medium 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 SSM within a few years in every environment. Similar to Figure 3, for each environment, we plot estimated field-level profits after T-year experiments.



Constructing scenario B

To generate a state vector $z = (z^1, z^2, z^3, z^4)$ for each site, we need to choose which site s and what values for $(z_s^1, z_s^2, z_s^3, z_s^4)$. For both purposes, we use random walk. Start from the mid site $(9, 5)$ with the initial values $(288, 3.2, 10, 10)$ assigned. Then, with probability of $1/3$, randomly either move right, move left, or stay. Similarly, with probability of $1/3$, randomly either move up, move down, or stay. This gives us the next site to consider. If the move means hitting a boundary, it stays at the site. Once moving into the new site, see if the site has already been assigned a state vector. If not, with probability of $1/3$, randomly perturb the state vector at the originating site by either -5%, 0%, or 5%. Continue the process until all sites are assigned a state vector.

Gaussian process

Gaussian process is a Bayesian nonparametric model, and its behavior is largely dependent on a choice of kernel and its hyperparameters [81]. 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, Larochelle, and Adams [82] and Stein [83], 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 $\nu \in \{1/2, 3/2, \infty\}$.

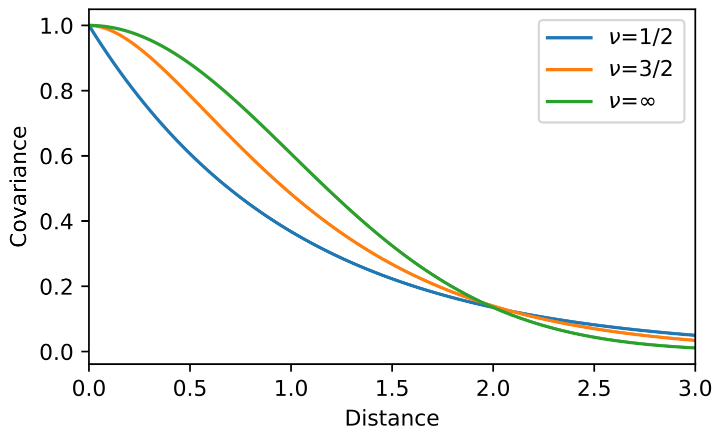


Figure 11: Matérn$_\nu$ kernels

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. $y^* = \max\{y_1, \ldots, y_{t-1}\}$. 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 [62, 63]:

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

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