A Conservative Stochastic Contextual Bandit Based Framework for Farming Recommender Systems

Abstract

The goal of this work is to develop a recommendation system for smart farming such that based on the details of the farm and the farming conditions, including information on the weather and soil properties, the system presents recommendations on the choice of the crop/seed, the type of fertilizers, and the amount of water for irrigation in order to maximize the overall net profit of the farmer. We refer to this problem as the crop-fertilizer-irrigation (CFI) recommender problem. In this paper, we propose a conservative, stochastic, contextual bandit formulation for solving the CFI problem, where the context captures the farm ID, weather indices and soil properties, the action set is the possible types of crops/seed, the types of fertilizers and irrigation, and the reward is the net profit of the farmer. Our bandit formulation is a conservative bandit setting since we incorporate constraints on the learned policy such that the learned policy need to satisfy certain performance criteria imposed by the farmer while maximizing the reward. Furthermore, our bandit formulation is also stochastic in the sense that the contexts are not observable, rather a distribution of the contexts are known. The stochastic bandit setting captures the uncertainty associated with the measurements of the weather indices and the soil properties. Based on the optimism in the face of uncertainty principle, we propose an algorithm to solve the bandit formulation of the CFI problem. To validate the performance of our approach we used the maize data collected through the G2F initiative. While we present initial model fitting results, the implementation and validation of the proposed algorithm are part of our future work.

Introduction

Making agricultural production system decisions is challenging due to risks and uncertainties associated with climate change, extreme weather events, high volatility in the crop price market, and increased regulatory pressure on the environment and public health (Hardaker et al. 2015). At the pre-season stage, a farmer must make crop selection decisions based on forecast/simulated weather, soil characteristics, available resources, expected prices of supplies such as water and fertilizer, and anticipated market value of the crop, all while adhering to budget and public regulations. Under uncertain conditions, farmers' crop choices – which crops to sow and how much acreage to devote to various crops – are crucial for achieving the best trade-off between profit and environmental footprint/regulation. The majority of farmers rely their seed selections on knowledge gathered from years of experience, the outcomes of prior decisions, and contacts with colleagues and consultants. Which, as it turns out, is not the optimal strategy in the face of uncertainty. Farmers require decision support tools to make a more informed decision.

Currently, available pre-season decision support tools are based on mechanistic crop models and data-driven models. Crop models including DSSAT (Jones et al. 1998), APSIM (Keating et al. 2003), and WOFOST (Eitzinger et al. 2004), are used to simulate crop yield under all potential what-if scenarios while accounting for uncertainty and facilitate crop selection. However, the model requires calibration for each field and crop type, which is costly, time consuming, and laborious (Lamsal et al. 2017). Data driven model includes linear models and its variants (Li et al. 2017; Iizumi et al. 2018; Drummond et al. 2003; Gonzalez-Sanchez, Frausto-Solis, and Ojeda-Bustamante 2014), nonlinear models such as support vector regression, Gaussian process regression (Shirley et al. 2020), decision tree (Romero et al. 2013), random forest (Jeong et al. 2016), neural networks (Kaul, Hill, and Walthall 2005; Abbasi-Yadkori, Pál, and Szepesvári 2011; Barbosa et al. 2020; Cunha, Silva, and Netto 2018). The generalizability of these models is dependent on data variety, quantity, and contexts. Uncertainty and risk were incorporated to these models by stochastic programming (Li and Hu 2020), portfolio optimization (Marko et al. 2017), and stochastic contextual bandit (Kirschner and Krause 2019). The stochastic contextual bandit formulation can account for weather, soil, and other variables' uncertainty but not for constraints such as budget limits. To account for the limitations, we extend the formulation to a conservative stochastic contextual bandit.

Problem Formulation: Conservative, Stochastic Bandit Formulation for Farming Recommendation

Consider a farmland denoted by $\mathcal{F}$, which is divided into an $n \times m$ grid, i.e., nm number of cells. Each cell in the grid is associated with a feature vector that captures the soil properties of the cell, including soil fertility levels and soil texture. The entire land $\mathcal{F}$ is associated with a feature vector that captures the climate indices such as heat degree days, frost days, highest temperature, lowest temperature, average temperature, and rainfall aggregated over specific periods. Let $\mathcal{X}$ be the set defined as $\mathcal{X} := \{\mathcal{X}_S\} \times \{\mathcal{X}_F\} \times \{\mathcal{X}_I\}$, where $\mathcal{X}_S$ is the set of all possible crops or seed types that can be grown in $\mathcal{F}$, $\mathcal{X}_F$ is the set of all available fertilizers that can be used in $\mathcal{F}$, and $\mathcal{X}_I$ is the amount of irrigation for $\mathcal{F}$. Additionally, often farmers impose performance constraints such as the net profit must be at least a certain value, and sustainability constraints such as limits on the amount of irrigation water and fertilizers and choices of fertilizers. For a given farmland and set of soil properties and climate indices, our goal is to provide recommendations for the crop/seed type, the fertilizer type, and the irrigation for each cell in the grid such that the annual net profit of the farmer is maximized and the associated constraints are satisfied. Here, farmer's annual net profit is defined as the total revenue subtracted by total management costs (i.e., cost of fertlizers and irrigation) during the farming process. We refer to this problem as the *crop-fertilizer-irrigation (CFI) recommender* problem.

In this work, we propose a contextual bandit formulation for the CFI recommender problem. Consider a context set $\mathcal{C} = \{c_1, c_2, \ldots\}$ of finite dimension. Each context $c_i \in \mathcal{C}$ is a d-dimensional vector that captures the soil properties and the weather indices. To consider the uncertainty associated with the weather and soil data, we consider a setting where the nature provides a distribution over the context set rather than the context itself, i.e., a prediction on the weather and soil data. At round t, the nature chooses a distribution $\mu_t \in \mathcal{P}(\mathcal{C})$ over the context set and samples a context realization $c_t \sim \mu_t$. The learner observes only μ_t and not c_t and chooses an action, say x_t. The unknown reward function is given by $f : \mathcal{C} \times \mathcal{X} \to \mathbb{R}$ such that for a given context $c_t \in \mathcal{C}$ and choice of crop or seed type x_S, fertilizer x_F, and irrigation x_I at time t, $f(x_t, c_t)$ is the resulting net profit, where $x_t = (x_S, x_F, x_I)$ and $x_t \in \mathcal{X}$. In addition, there exists a baseline (farmer's) policy π_b that at each round t, selects action $b_t \in \mathcal{X}$ and incurs the expected reward $z(b_t, c_t)$.

At round t, the learner observes μ_t and makes a decision x_t, i.e., recommends a crop or seed type, fertilizer, and irrigation, based on a prediction on soil properties and weather conditions. Then the nature provides a reward $y_t = f(x_t, c_t) + \epsilon_t$, where $y_t \in \mathbb{R}$, ϵ_t is a $\sigma-$subGaussian, additive noise and satisfies Assumption 1. In addition, at each round t the learner queries the farmer to present his/her policy b_t based on the prior experience and preferences. The expected rewards of the actions taken by the farmer is denoted as $z(b_t, c_t)$. We assume that $z(b_t, c_t)$ is known since we have access to a large amount of data generated by the baseline policy, i.e., farmer's strategy, and thus have a good estimate of its performance. In order to solve the CFI recommender problem, our aim is to learn an optimal mapping/policy $g : \mathcal{C} \to \mathcal{X}$ of contexts (soil & weather conditions) to the crops, fertilizer, and irrigation such that the cumulative reward, $\sum_{t=1}^{T} f(x_t, c_t)$ is maximized while simultaneously satisfying the constraints imposed by the farmer. The constraints are such that at round t, the difference between the performances of the baseline and the learner's policies should remain above a pre-defined fraction $\alpha \in (0, 1)$ of the baseline performance. Formally, our aim is to minimize the cumulative regret

$$\mathcal{R}_T = \sum_{t=1}^{T} \Big(f(x_t^\star, c_t) - f(x_t, c_t) \Big), \tag{1}$$

such that

$$\sum_{i=1}^{t} f(x_t, c_t) \geq (1 - \alpha) \sum_{i=1}^{t} z(b_t, c_t) \tag{2}$$

at each round, where $x^\star := \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim \mu_t}[f(x, c)]$ is the best choice of crop/seed, fertilizer, and irrigation for the given context μ_t, T is the number of rounds, and $\alpha \in (0, 1)$ is the maximum decrease in the performance the decision maker is willing to accept. The constraint metric ensures that the learned policies satisfy certain requirements of the farmer such as type and amount of fertilizer and amount of water, and net profit. We focus on linearly parameterized reward functions $f(x, c) = \phi_{x,c}^\top \theta$ with given feature vectors $\phi_{x,c} \in \mathbb{R}^d$ for $x \in \mathcal{X}$ and $c \in \mathcal{C}$, and an unknown reward parameter vector $\theta \in \mathbb{R}^d$. Thus the reward at round t is denoted as

$$y_t = \phi_{x_t, c_t}^\top \theta + \epsilon_t, \tag{3}$$

where θ and ϕ satisfy Assumption 2. A bandit setting with a linear reward structure is referred to as *linear bandit* setting (Bubeck and Cesa-Bianchi 2012).

Assumptions 1 and 2 are presented below.

Assumption 1 *Each element ϵ_t of the noise sequence $\{\epsilon_t\}_{t=1}^{\infty}$ is conditionally $\sigma-$subGaussian, i.e.,*

$$\textit{For all } \zeta \in \mathcal{R}, \mathbb{E}[e^{\zeta \epsilon_t} | x_{1:t, \epsilon_{1:t-1}}] \geq exp(\frac{\zeta^2 \sigma^2}{2}).$$

Assumption 2 *There exists constant $A, D \geq 0$ such that $\|\theta\|_2 \leq A$, $\|\phi_{x,c_t}\|_2 \leq D$, and $\phi_{x,c_t}^\top \theta \in [0, 1]$, for all t and all $x \in \mathcal{X}$.*

Background: Linear Contextual Bandits

In this section, we briefly review the results from the linear contextual bandit literature. The linear contextual bandit problem has been extensively studied and different solution approaches have been proposed (Li et al. 2010; Allesiardo, Féraud, and Bouneffouf 2014; Agrawal and Goyal 2013; Abbasi-Yadkori, Pál, and Szepesvári 2011). In the linear contextual bandit setting, there are no constraints that need to be satisfied by the learner and the context in round t is known and hence it is a special case of the bandit setting considered in this paper with no constraints and the choice of the distribution μ_t as a Dirac delta distribution denoted as $\mu_t = \delta_{c_t}$ for all t. In round t of the linear contextual bandit setting, the nature presents an action-context feature vector $\Psi_t = \{\phi_{x,c_t} : x \in \mathcal{X}\} \subset \mathbb{R}^d$. The learner then selects an action $x_t \in \mathcal{X}$ and observes a noisy reward $y_t = \phi_t^\top \theta + \epsilon_t$, where ϵ_t is conditionally σ-subGaussian. The goal here is to minimize the cumulative regret $\mathcal{R}_T = \sum_{t=1}^{T} \phi_t^{\star\top} \theta - \phi_t^\top \theta$, where $\phi_t^\star = \arg\max_{\phi \in \Psi_t} \phi^\top \theta$ is the feature vector corresponding to the best action in round t.

A linear contextual bandit setting with uncertainty in the context is studied in (Lamprier, Gisselbrecht, and Gallinari 2018; Kirschner and Krause 2019), specifically (Kirschner and Krause 2019) considered a setting where the context itself is not observable rather a distribution on the context is available. Often decision making problems are associated with *safety or performance* constraints and the objective of the learner is to learn a policy that maximizes the cumulative reward and guarantees that the learned policy performs at least as well as a baseline. Constrained linear contextual bandits are studied in (Kazerouni et al. 2017; Amani, Alizadeh, and Thrampoulidis 2019; Russo and Van Roy 2014; Daulton et al. 2019). In this work we built on the work in (Kirschner and Krause 2019; Kazerouni et al. 2017) to address the CFI-recommender problem in which the contexts are uncertain (predictions of weather and soil data) and the farmland is subject to certain constraints.

Solution Approach: Conservative Stochastic Contextual Bandit

In this section, we present the algorithm for solving the CFI-recommender problem. We observe the predictions on the weather and the soil properties, μ_t, for each grid in the farmland instead of the accurate measurements c_t, and the features ϕ_{x,c_t}. Given the distribution μ_t, we construct the expected feature vector, $\Psi_t = \{\bar{\psi}_{x,\mu_t} : x \in \mathcal{X}\}$ where $\{\bar{\psi}_{x,\mu_t} := \mathbb{E}_{c\sim\mu_t}[\phi_{x,c}]\}$ (step: 4) (Kirschner and Krause 2019). We note that, each feature $\bar{\psi}_{x,\mu_t}$ corresponds to exactly one action $x \in \mathcal{X}$ and we use Ψ_t as the feature context set at time t. The proposed algorithm is based on the *optimism in the face of uncertainty* principle, where the algorithm maintains a confidence set $\mathcal{B}_t \subset \mathcal{R}^d$ that contains the unknown parameter vector θ with high probability (Abbasi-Yadkori, Pál, and Szepesvári 2011). The algorithm then chooses an optimistic estimate $\tilde{\theta}_t = \arg\max_{\hat{\theta}\in\mathcal{B}_t} (\max_{x\in\mathcal{X}} \bar{\psi}_{x,\mu_t}^\top \hat{\theta})$ and chooses an action $x'_t = \arg\max_{x\in\mathcal{X}} \bar{\psi}_{x,\mu_t}^\top \tilde{\theta}_t$. Equivalently the algorithm chooses the pair $(x'_t, \tilde{\theta}_t) \in \arg\max_{(x,\hat{\theta})\in\mathcal{X}\times\mathcal{B}_t} \bar{\psi}_{x,\mu_t}^\top \hat{\theta}$ which jointly maximizes the reward.

To ensure that the action chosen by the algorithm guarantees satisfaction of the constraints imposed by the farmer, the algorithm plays the action x'_t only if it satisfies the constraint for the worst choice of the parameter $\hat{\theta} \in \mathcal{B}_t$ (Kazerouni et al. 2017). We formally define this by introducing two sets S_{t-1}^b and S_{t-1}. Let S_{t-1} be the set of rounds i before round t at which the algorithm has played the optimistic action, i.e., $x_i = x'_i$. Then $S_{t-1}^b = \{1, 2, \ldots, t-1\} - S_{t-1}$ is the set of rounds j before round t at which the algorithm has followed the baseline policy, i.e., $x_j = b_j$. To ensure that constraint in Eq. (2) is satisfied the algorithm plays optimal action $x_t = x'_t$ at round t if it satisfies

---

Algorithm 1: Pseudocode for Conservative Stochastic Contextual Bandit

Input: $\alpha, \mathcal{B}, \mathcal{F}$
Initialize: $S_0 = \emptyset, \ell_0 = 0 \in \mathbb{R}^d, \mathcal{B}_1 = \mathcal{B}$

1: **for** $t = 1, 2, \ldots, T$ **do**
2: Nature chooses $\mu_t \in \mathcal{P}(\mathcal{C})$
3: Learner observes μ_t
4: Set $\Psi_t = \{\bar{\psi}_{x,\mu_t} : x \in \mathcal{X}\}$ where $\{\bar{\psi}_{x,\mu_t} := \mathbb{E}_{c\sim\mu_t}[\phi_{x,c}]\}$
5: Query baseline (farmer's) strategy $b_t \leftarrow \pi_f(\Psi_t)$
6: Find $(x'_t, \tilde{\theta}_t) \in \arg\max_{(x,\hat{\theta})\in\mathcal{X}\times\mathcal{B}_t} \bar{\psi}_{x,\mu_t}^\top \hat{\theta}$
7: Compute $L_t = \min_{\hat{\theta}\in\mathcal{B}_t} \langle \ell_{t-1} + \bar{\psi}_{x'_t,\mu_t}, \hat{\theta}\rangle$
8: **if** $L_t + \sum_{i\in S_{t-1}^b} z(b_t, c_t) \geq (1-\alpha)\sum_{i=1}^t z(b_t, c_t)$ **then**
9: Play $x_t = x'_t$ and observe reward y_t in Eq. (3)
10: Set $\ell_t = \ell_{t-1} + \bar{\psi}_{x_t,\mu_t}$, $S_t = S_{t-1} \cup t$, $S_t^b = S_{t-1}^b$
11: Given x_t, y_t construct $\mathcal{B}_{t+1}$ using Eq. (5)
12: **else**
13: Play $x_t = b_t$ and observe reward y_t in Eq. (3)
14: Set $\ell_t = \ell_{t-1}$, $S_t = S_{t-1}$, $S_t^b = S_{t-1}^b \cup t$, $\mathcal{B}_{t+1} = \mathcal{B}_t$
15: **end if**
16: **end for**

$$\min_{\hat{\theta}\in\mathcal{B}_t} \Big[\sum_{i\in S_{t-1}^b} z(b_t, c_t) + (\sum_{i\in S_{t-1}} \bar{\psi}_{x_i,\mu_i})^\top \hat{\theta} + \bar{\psi}_{x'_t,\mu_t}^\top \hat{\theta} \Big]$$

$$\geq (1-\alpha) \sum_{i=1}^t z(b_i, c_i),$$

and plays the action chosen by the farmer, i.e., $x_t = b_t$ otherwise. For details on the construction of the confidence set $\mathcal{B}_t$, we refer to Section 3.1 in (Kazerouni et al. 2017).

Construction of the Confidence Set $\mathcal{B}_t$: We denote the confidence set in round t as $\mathcal{B}_t$. The proposed algorithm starts by the most general confidence set i.e., $\mathcal{B}_1 = \mathcal{B}$, and updates the confidence set only when the optimistic action proposed by the learner is played. This is because that unless the learner's action is played, no additional information is gained about the unknown parameter θ. Let $S_t = \{i_1, i_2, \ldots, i_{m_t}\}$ be the set of rounds up to and including t during which the the algorithm played the optimistic action. Here $m_t = |S_t|$. For a fixed value $\lambda > 0$, the regularized least square estimate of $\hat{\theta}$ at round t is given by

$$\bar{\theta}_t = \left(\Phi_t \Phi_t^\top + \lambda\right)^{-1} \Phi_t Y_t, \tag{4}$$

where $\Phi_t = [\bar{\psi}_{x_{i_1},\mu_{i_1}}, \bar{\psi}_{x_{i_2},\mu_{i_2}}, \ldots, \bar{\psi}_{x_{m_t},\mu_{m_t}}]$ and $Y_t = [y_{i_1}, y_{i_2}, \ldots, y_{m_t}]^\top$. For a given confidence parameter $\delta \in (0,1)$, we construct the confidence set for the next round $t+1$ as

$$\mathcal{B}_{t+1} = \{\hat{\theta} \in \mathcal{R}^d : \left\|\hat{\theta} - \bar{\theta}_t\right\|_{V_t} \leq \beta_{t+1}\}, \tag{5}$$

where $\beta_{t+1} = \sigma\sqrt{d\log(\frac{1+(m_t+1)D^2/\lambda}{\delta})} + \sqrt{\lambda}A$, $V_t = \lambda I + \Phi_t\Phi_t^\top$, and the weighted norm is defined as $\|u\|_V = \sqrt{u^\top V u}$ for any $u \in \mathcal{R}^d$ and positive definite $V \in \mathcal{R}^{d\times d}$.

Experimental Analysis

Crop Data

We use a maize yield data set acquired over four years by the maize genomes to fields (G2F) initiative (McFarland et al. 2020), a multi-institutional effort in North America over 68 unique locations. The data set includes yields, planting dates, flowering times, and harvest dates, as well as hourly weather data from in-field weather stations, such as temperature, humidity, solar radiation, rainfall, and soil wind speed, as well as soil characteristics such as soil texture, organic matter, texture, and nitrogen, phosphorous, potassium, sulfur, and sodium levels (in parts per million). There are 2158 yield measurements for 24 crops collected from 22 different locations in this data set.

For this experiment, the weather data of the whole growing season was summarized by crop growth stages as in (Holzkämper, Calanca, and Fuhrer 2013). These are average daily solar radiation [MJ/m2], average daily minimum temperature below 0 ^{0}C in absolute values [^{0}C] as a measure of frost impacts, average daily mean temperature [^{0}C] as a measure of temperature determining plant growth, average daily maximum temperature above 35 C [^{0}C] as a measure of heat stress, and average photoperiod.

Experiments

We first constructed a data set $\mathcal{D} = \{(c_i, x_i, y_i)\}$, where for each data point $i = 1, 2, \ldots, 2158$, $c_i \in \mathbb{R}^{28}$ is a $28-$dimensional vector that includes 6-dimensional weather and soil data information (% of sand, % of silt, % of clay in the soil, daily average temperature, radiation, and photosynthesis) and a $22-$dimensional one-hot encoding that captures the field ID, and x_i, y_i are the seed/crop identifier, yield, respectively. The set $\mathcal{D}$ is of size 2158 which is the size of our data set. As an initial step before implementing our proposed algorithm, we first fit a bilinear model (Koren, Bell, and Volinsky 2009) such that $y_i \approx c_i^\top W V_{x_i}$, where $V_{x_i} \in \mathbb{R}^{10}$ is the feature vector for crop type x_i (Kirschner and Krause 2019). Our data set consists of 24 varieties of crops and hence there are 24 feature vectors, $V_1, V_2, \ldots, V_{24}$. The bilinear model captures the correlation between site features $c_i^\top W$ and V_{x_i} for each data point and serves as the interactive setting that provides the rewards (yield) for our bandit setting.

We fitted a bilinear model on the historical maize data, collected through the G2F initiative (McFarland et al. 2020), via stochastic gradient descent using the loss function below (Koren, Bell, and Volinsky 2009)

$$L(V, W) = \sum_{i=1}^{n}(y_i - c_i^T W V_{x_i})^2 + \lambda_v||V_{x_i}||^2 + \lambda_w||W||^2,$$

where λ_v and λ_w denotes the regularization terms. Training this model for 300 iterations resulted in a mean square error loss of 0.002 using a learning rate of 0.015, $\lambda_v = \lambda_w = 0.001$ and a latent dimension of 10, i.e., $V_{x_i} \in \mathbb{R}^{10}$ for all i. Figure 1 illustrates the model predicted yield values versus the true yield values in the data set. As seen in the figure, the yield values predicted by the model correlates well with the true yield values, signifying that the trained bi-linear model is a good candidate to test our proposed bandit algorithm. Our future work will include validating this crop model on additional data, including additional relevant context, and implementing our proposed conservative stochastic contextual bandit algorithm, Algorithm 1, using this model.

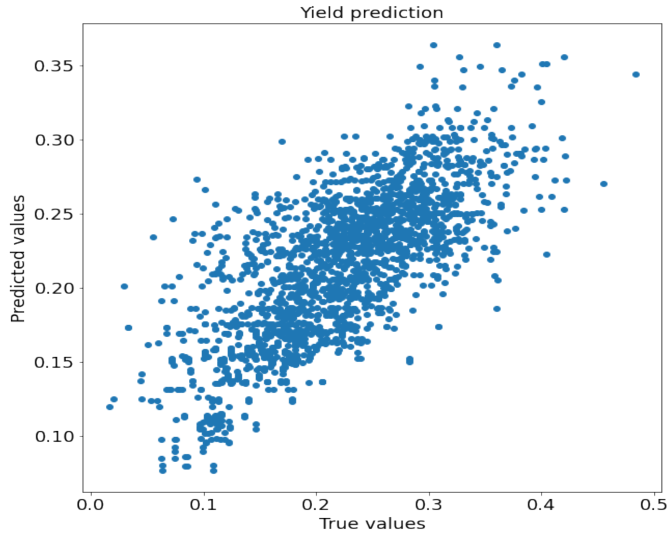

Figure 1: Visualization of yield values predicted by the model versus true yield values. The predicted values shows correlation with the true values.

Conclusion and Future Work

In this paper, we presented a conservative contextual bandit-based recommendation system for farming such that given the details of the farm, the climatic conditions, and soil properties, the system provides recommendation on the type of crops/seeds, fertilizers, and amount of water for irrigation, in order to maximize the net profit of the farmer. In our formulation, we incorporated the domain knowledge and experience of the farmer into the decision making process by including performance constraints such that the farmer can impose certain conditions on the minimum expected yield, type of the fertilizers and so on. Including these constraints into the model gives the farmer a certain amount of control on the farming decisions instead of solely relying on the decisions recommended by the learner. The uncertainties associated with the weather and soil property measurements are captured by considering a stochastic contextual bandit where the actual contexts are unknown rather only the predictions are available. To validate the performance of our approach we plan to use the maize data collected through the G2F initiative (McFarland et al. 2020). While some initial model fitting is completed, the implementation and validation of the proposed algorithm are part of our future work. As part of our future work, we also plan to investigate the convergence of the proposed algorithm.

References

Abbasi-Yadkori, Y.; Pál, D.; and Szepesvári, C. 2011. Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems*, 24: 2312–2320.

Agrawal, S.; and Goyal, N. 2013. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, 127–135.

Allesiardo, R.; Féraud, R.; and Bouneffouf, D. 2014. A neural networks committee for the contextual bandit problem. In *International Conference on Neural Information Processing*, 374–381.

Amani, S.; Alizadeh, M.; and Thrampoulidis, C. 2019. Linear stochastic bandits under safety constraints. *arXiv:1908.05814*.

Barbosa, A.; Trevisan, R.; Hovakimyan, N.; and Martin, N. F. 2020. Modeling yield response to crop management using convolutional neural networks. *Computers and Electronics in Agriculture*, 170: 105197.

Bubeck, S.; and Cesa-Bianchi, N. 2012. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv:1204.5721*.

Cunha, R. L.; Silva, B.; and Netto, M. A. 2018. A scalable machine learning system for pre-season agriculture yield forecast. In *2018 IEEE 14th International Conference on e-Science (e-Science)*, 423–430. IEEE.

Daulton, S.; Singh, S.; Avadhanula, V.; Dimmery, D.; and Bakshy, E. 2019. Thompson sampling for contextual bandit problems with auxiliary safety constraints. *arXiv:1911.00638*.

Drummond, S. T.; Sudduth, K. A.; Joshi, A.; Birrell, S. J.; and Kitchen, N. R. 2003. Statistical and neural methods for site–specific yield prediction. *Transactions of the ASAE*, 46(1): 5.

Eitzinger, J.; Trnka, M.; Hösch, J.; Žalud, Z.; and Dubrovskỳ, M. 2004. Comparison of CERES, WOFOST and SWAP models in simulating soil water content during growing season under different soil conditions. *Ecological modelling*, 171(3): 223–246.

Gonzalez-Sanchez, A.; Frausto-Solis, J.; and Ojeda-Bustamante, W. 2014. Attribute selection impact on linear and nonlinear regression models for crop yield prediction. *The Scientific World Journal*, 2014.

Hardaker, J. B.; Lien, G.; Anderson, J. R.; and Huirne, R. B. 2015. *Coping with risk in agriculture: Applied decision analysis*. Cabi.

Holzkämper, A.; Calanca, P.; and Fuhrer, J. 2013. Identifying climatic limitations to grain maize yield potentials using a suitability evaluation approach. *Agricultural and Forest Meteorology*, 168: 149–159.

Iizumi, T.; Shin, Y.; Kim, W.; Kim, M.; and Choi, J. 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Climate Services*, 11: 13–23.

Jeong, J. H.; Resop, J. P.; Mueller, N. D.; Fleisher, D. H.; Yun, K.; Butler, E. E.; Timlin, D. J.; Shim, K.-M.; Gerber, J. S.; Reddy, V. R.; et al. 2016. Random forests for global and regional crop yield predictions. *PLoS One*, 11(6): e0156571.

Jones, J.; Tsuji, G.; Hoogenboom, G.; Hunt, L.; Thornton, P.; Wilkens, P.; Imamura, D.; Bowen, W.; and Singh, U. 1998. Decision support system for agrotechnology transfer: DSSAT v3. In *Understanding options for agricultural production*, 157–177. Springer.

Kaul, M.; Hill, R. L.; and Walthall, C. 2005. Artificial neural networks for corn and soybean yield prediction. *Agricultural Systems*, 85(1): 1–18.

Kazerouni, A.; Ghavamzadeh, M.; Abbasi-Yadkori, Y.; and Van Roy, B. 2017. Conservative contextual linear bandits. *Advances in Neural Information Processing Systems*.

Keating, B. A.; Carberry, P. S.; Hammer, G. L.; Probert, M. E.; Robertson, M. J.; Holzworth, D.; Huth, N. I.; Hargreaves, J. N.; Meinke, H.; Hochman, Z.; et al. 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3-4): 267–288.

Kirschner, J.; and Krause, A. 2019. Stochastic bandits with context distributions. *Advances in Neural Information Processing Systems*, 32: 14113–14122.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8): 30–37.

Lamprier, S.; Gisselbrecht, T.; and Gallinari, P. 2018. Profile-based bandit with unknown profiles. *The Journal of Machine Learning Research*, 19(1): 2060–2099.

Lamsal, A.; Welch, S.; Jones, J.; Boote, K.; Asebedo, A.; Crain, J.; Wang, X.; Boyer, W.; Giri, A.; Frink, E.; et al. 2017. Efficient crop model parameter estimation and site characterization using large breeding trial data sets. *Agricultural systems*, 157: 170–184.

Li, L.; Chu, W.; Langford, J.; and Schapire, R. E. 2010. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, 661–670.

Li, Q.; and Hu, G. 2020. Multistage stochastic programming modeling for farmland irrigation management under uncertainty. *Plos one*, 15(6): e0233723.

Li, Q.; Hu, G.; Jubery, T. Z.; and Ganapathysubramanian, B. 2017. A farm-level precision land management framework based on integer programming. *PloS one*, 12(3): e0174680.

Marko, O.; Brdar, S.; Panić, M.; Šašić, I.; Despotović, D.; Knežević, M.; and Crnojević, V. 2017. Portfolio optimization for seed selection in diverse weather scenarios. *PloS one*, 12(9): e0184198.

McFarland, B. A.; AlKhalifah, N.; Bohn, M.; Bubert, J.; Buckler, E. S.; Ciampitti, I.; Edwards, J.; Ertl, D.; Gage, J. L.; Falcon, C. M.; et al. 2020. Maize genomes to fields (G2F): 2014–2017 field seasons: genotype, phenotype, climatic, soil, and inbred ear image datasets. *BMC research notes*, 13(1): 1–6.

Romero, J. R.; Roncallo, P. F.; Akkiraju, P. C.; Ponzoni, I.; Echenique, V. C.; and Carballido, J. A. 2013. Using classification algorithms for predicting durum wheat yield in the province of Buenos Aires. *Computers and electronics in agriculture*, 96: 173–179.

Russo, D.; and Van Roy, B. 2014. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4): 1221–1243.

Shirley, R.; Pope, E.; Bartlett, M.; Oliver, S.; Quadrianto, N.; Hurley, P.; Duivenvoorden, S.; Rooney, P.; Barrett, A. B.; Kent, C.; et al. 2020. An empirical, Bayesian approach to modelling crop yield: Maize in USA. *Environmental Research Communications*, 2(2): 025002.