# Understanding the farm data lifecycle: collection, use, and impact of farm data on U.S. commercial corn and soybean farms

Nathanael M. Thompson[1] · Nathan D. DeLay[1] · James R. Mintert[1]

## Abstract

Enthusiasm regarding the "digital agriculture" revolution is widespread, yet objective research on how commercial farms actually use data and data services remains limited. The purpose of this research is to better understand the current positioning of U.S. commercial corn and soybean farms within the farm data lifecycle, including the collection, use, and impact of farm data. Using survey data from a sample of 800 commercial-scale U.S. corn and soybean farms, the factors associated with progression within the farm data lifecycle are examined. Results indicate that the majority of commercial U.S. corn and soybean farms collect data, indicate that the data they collect influences their decisions, and perceive positive yield benefits as a result of their data-informed decisions. However, farms vary in intensity of their data usage. Investments in data management and analysis resources are associated with progression within the farm data lifecycle. These investments comprise software products that manage and analyze data, including creating GPS maps, layering different data sources, and generating recommendations. Investments in human capital, either in on-farm employees with designated data responsibilities or in trusted off-farm service providers, are also associated with progression within the farm data lifecycle. Farms that have not yet invested in these types of data management and data analysis resources may be forfeiting the potential benefits associated with using their farm's data to improve on-farm decision making.

**Keywords** Commercial farms · Corn and soybeans · Digital agriculture · Farm data lifecycle · Precision agriculture

## Introduction

A great deal of enthusiasm over "digital agriculture" and "big data" in agriculture has emerged among industry, venture capital, and an eager farm press. Several large-scale acquisitions suggest the value of farm data to be vast. Monsanto described their purchase

✉ Nathanael M. Thompson
   thomp530@purdue.edu

[1] Department of Agricultural Economics, Purdue University, West Lafayette, IN 47907, USA

of The Climate Corporation for over $900 million in 2013 as their "entry ticket into a $20 billion market opportunity," while John Deere believed precision agriculture could help increase their company's value by $25 billion (Plume, 2014). Similar enthusiasm has also permeated the academic literature in recent years (e.g., Coble et al., 2018; Weersink et al., 2018; Woodard et al., 2018).

Despite widespread interest, there is a noticeable deficit of objective research on how commercial farms actually use data and data services. For example, what kinds of data are farmers actually collecting? What software platforms do producers subscribe to and what features do these software solutions offer? Do producers share their data with outside service providers? Most importantly, how do these data and software services integrate into existing production systems to impact decision making and, subsequently, farm outcomes (e.g., yield, efficiency, profit)? There is a clear need to better understand how commercial farms in the United States use the data they collect and the value that data brings to their operations relative to the technical, and sometimes impractical, offerings of data service providers. Researchers and extension educators can better serve agricultural producers, service providers, and developers by addressing these questions.

Previous research has largely focused on adoption of precision agriculture hardware (e.g., Schimmelpfennig, 2016; Zhou et al., 2017; Lowenberg-Deboer & Erickson, 2019). Hardware adoption is an important component of a digital agriculture system and is a necessary step for implementing advanced data-intensive technologies (Khanna et al., 1999). However, hardware adoption by itself does not directly measure farmers' use of the data they collect. Some studies have implicitly sought to quantify data use by evaluating the adoption of precision agriculture technology bundles (Lambert et al., 2015; Schimmelpfennig & Ebel, 2016; Griffin et al., 2017; Miller et al., 2017, 2019). That is, combinations of precision agriculture technologies likely serve as a proxy for the transition from data collection to data use (e.g., yield monitor and variable rate technology [VRT]). While evaluating the adoption of technology bundles lends some insight into the progression from data collection to data use, it still focuses on hardware adoption, and only implies data use.

Schimmelpfennig (2016) characterizes this relationship between data collection and data use through the concept of "information flows." For example, producers must first collect geo-referenced yield data using a yield monitor; that information is then used to create a yield map using farm software or a service provider; the yield map is then used to create prescriptions for VRT input applications, such as seed or fertilizer. Using USDA Agricultural Resource Management Survey (ARMS) data, Schimmelpfennig (2016) shows that in 2010 48% (70%) of U.S. corn farms (acres) owned a yield monitor, but just 25% (44%) made a geo-referenced yield map and only 19% (28%) applied inputs using VRT. It is important to point out that yield monitor data can be used for purposes other than making yield maps and VRT prescriptions. For example, USDA ARMS data indicate that in 2010 U.S. corn farmers used yield monitor data to monitor crop moisture (52% of corn acres), conduct on-farm experiments (20%), and document yields (28%) (USDA Economic Research Service [ERS], 2020). In addition, in some cases yield monitor data may not indicate that VRT input applications are warranted. In any case, it is evident that not all farms collecting data are using it to make actionable decisions on their farms. The question that remains is to what extent has the gap between data collection and data use narrowed and what factors are influencing progress?

It is also important to point out that the 2010 ARMS data referenced here are the most recent ARMS data publicly available for corn via the ARMS Tailored Reports (USDA ERS, 2020). The 2016 ARMS survey is the most recent iteration of the ARMS survey for corn, but data from this survey are not yet publicly available. Lowenberg-DeBoer and

Erickson (2019) summarize some of the general precision agriculture adoption rates for the 2016 ARMS corn survey.

The current research seeks to move beyond the study of hardware adoption towards better understanding of the farm data lifecycle. Instead of focusing on the adoption of hardware, this study employs a survey instrument where respondents are explicitly asked what types of data they collect (i.e., data collection), the degree to which the data they collect influences their farm management decisions (i.e., data use), and finally, their perception of the outcomes from their data-informed decisions (i.e., data impact). The objective of this study is to identify where commercial U.S. corn and soybean farms lie in the farm data lifecycle and to determine the factors associated with producers' progress in the collection, use, and impact of farm data on their operations. The research hypothesis motivating the current study is that data collection is common among commercial farms, but the extent to which farm data is used to make actionable decisions lags behind data collection. It is also hypothesized that farm/farmer demographics, data management and analysis resources, technology use, and data sharing will be associated with progression within the various stages of the farm data lifecycle.

## Materials and methods

A data lifecycle provides a high-level framework for representing the stages of data throughout their life (Demestichas & Daskalakis, 2020). Adapting and extending Schimmelpfennig's (2016) concept of "information flows," the farm data lifecycle described here is simplified into a three-stage process (Fig. 1):

1. Stage #1: Data collection—farms choose to collect data or not.
2. Stage #2: Data use—farms that collect data decide the extent to which their data will influence their decision making.
3. Stage #3: Data impact—farms whose data influences their decisions evaluate the impact of data-informed decisions on farm output, efficiency, profitability, etc.

In practice, there are a variety of farm data types that can be collected in stage #1. For simplicity, this research focuses on three types of data: yield data, grid or zone soil sample data, and aerial or satellite imagery data. It is important to point out that different farm data types may have different lifespans. For example, soil organic matter (SOM) changes slowly. Therefore, SOM data may be collected once and that data may influence decisions for several years. Conversely, satellite/aerial imagery data may be used to detect and manage in-season crop deficiencies, but may be irrelevant to next year's management decisions. Further, other data types, such as yield monitor data, may be collected annually for several years and layered to enhance its decision value. While the issue of data lifespan and the role of data collection frequency in decision making is not taken up directly in this study, it is an important aspect of the farm data lifecycle to be explored in more detail in future research.

Similar to the various data types in stage #1, a number of decisions could be influenced by data in stage #2 of the farm data lifecycle. Respondents indicating they collected one or more of the queried types of data were asked in stage #2 about the extent to which that data influenced three different decisions: seeding rate decisions, nutrient management decisions, and drainage investment decisions. In practice, there is a
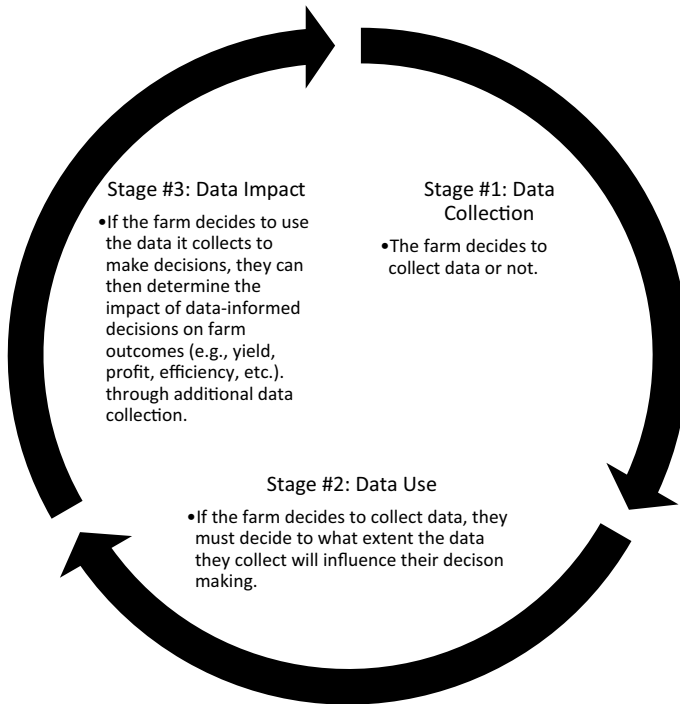
**Fig. 1** Farm data lifecycle

continuous scale of the degree to which farm data can influence decision making. However, for simplicity, respondents were given three potential levels of impact: not at all, somewhat, or a lot.

Finally, in stage #3 of the farm data lifecycle the impact of data-informed decisions on farm output was assessed for each of the three possible decisions in stage #2 (seeding rate, nutrient management, drainage investment). For example, if a respondent indicated the data they collected influenced their seeding rate decisions somewhat or a lot, they were asked whether their data-informed seeding rate decisions decreased, did not change, or increased their farm yield. Again, the range of yield impacts associated with data-informed decisions span a continuous scale, but discrete yield impacts (decrease, no change, or increase) were used to simplify the survey. Notice that the third and final stage of the farm data lifecycle described here does not necessarily imply that the data "die." Instead, stage #3 initiates the next iteration of the cycle as data impacts can only be evaluated through additional data collection (stage #1). Death of data relates to data lifespan, which was not examined in this research.

While a variety of measures could be used to evaluate the impact of data-informed decisions, yield was used in this study as it is the value for which respondents were expected to have the easiest recall. An economic measure, such as profit or efficiency, would be preferred. However, few farms have directly estimated the return on investment for precision agriculture technology or data investments (Pope & Sonka, 2020). Therefore, it was expected that asking farmers to recall or mentally account for the profit impact of their data-informed decisions would be laborious and subject to significant

measurement error. Conversely, farms are often conversant regarding yield impacts of their decisions and can easily recall these values. Nonetheless, yield impacts ascribed to data-informed decision making reported here likely represent respondent perceptions of yield impacts, as opposed to actual yield impacts, given the complexity associated with causally identifying the impact of these data-informed decisions on farm outcomes. Nonetheless, farmer perceptions are often what determine adoption (or disadoption) decisions (Pannell et al., 2006), and despite their limitations, are an important part of examining the current state of the farm data lifecycle.

In addition to questions regarding the farm data lifecycle, respondents were asked several questions regarding data management and sharing practices to gauge the farm's investment in a data strategy. Demographic questions were included in the survey to collect information about the farm.

A phone survey of U.S. commercial corn and soybean producers conducted from August 5, 2019, to August 30, 2019 was used to collect the data for this study. Institutional Review Board (IRB) approval was obtained for the study from Purdue University (IRB Protocol #1906022382). The survey list frame of commercial corn and soybean producers was purchased from *Farm Journal*, and the survey was administered by PRISM Marketing Group by phone. All enumerators read from the same survey script to ensure consistency. Given that data collection was done by phone, the survey was designed to be completed by respondents in less than 10 min. Questions were made to be short and easy to understand to encourage producer responses. Participation was voluntary and no compensation was provided for participating. A copy of the survey is available in the supplementary appendix (Appendix A).

The survey was intentionally targeted toward commercial U.S. corn and soybean farms defined as operations with farmland of 1000 acres or more. This targeted approach is taken given the propensity of large operations to collect these types of data. To ensure operation size diversity within the sample, quotas were imposed for survey sampling procedures. The USDA's 2017 Census of Agriculture reported 172,793 farms with more than 1000 acres of farmland operated in the United States (USDA NASS, 2020). Notice, this size classification is based on farmland acres operated which is distinct from cropland acres or corn and soybean acres. Given this population, a survey sample size of 383 is necessary to ensure a sample with a confidence level of 95% and a margin error of 5%. However, over half of farms with more than 1000 acres of farmland operated (87,666 farms) have less than 2000 acres of farmland operated (USDA NASS, 2020). To ensure that the sample was representative of larger-scale farms, and not just those operating less than 2000 acres of farmland, quotas were imposed on the data collection process that required the final sample to include at least 400 respondents farming between 1000 and 1999 acres of farmland and at least 400 responses farming 2000 acres or more.

The list frame was filtered based on farm size information available in the frame to help reach respondents that met the size criteria. Observations were randomly drawn from the filtered list frame, and farm size quotas were implemented using a screening question at the beginning of the survey, "How many total acres do you operate? Less than 1000 acres, 1000–1999 acres, 2000–4999 acres, or 5000 or more acres." In total, 7841 farmers were contacted that met the size criteria. Of those, 934 started the survey and 800 completed the survey for a response rate of 10% – 400 in the 1000–1999 acre category and 400 in the ≥ 2000 acre category.

Data are summarized and Fisher's (1922) exact test, a variation of the $\chi^2$ test, is used to determine if the responses across subsamples are statistically different. Significant relationships identified using Fisher's exact test represent associations between variables.

However, it is important to point out that these relationships do not imply causality. Rather, the Fisher (1922) approach can reveal useful correlations between a farm's stage in the data lifecycle and the farm data tools and practices employed by the farm, as well as key farm and operator characteristics.

## Results and discussion

Summary statistics for the full sample are reported in Table 1. Per sampling quotas, half of the farms surveyed operate between 1000 and 1999 acres, and half operate 2000 acres or more (36% are between 2000 and 4999 acres and 14% have 5000 acres or more). It is important to reiterate that the sample is intentionally not representative of all U.S. farms. Instead, the sample is specifically focused on commercial-scale farms (defined as farming 1000 acres or more) given their propensity to be involved in digital agriculture and this study's objective to learn more about the gap between data collection and data use and factors influencing progress in narrowing the gap between the two. More than 80% of

**Table 1** Summary statistics ($n=800$)

| Variable | Percent of respondents |
| --- | --- |
| *Farm/farmer demographics* | |
| Farm size | |
|   1000–1999 acres | 50% |
|   2000–4999 acres | 36% |
|   $\geq 5000$ acres | 14% |
| Primary farm owner/operator age | |
|   $<20$ years | $<1\%$ |
|   20–35 years | 2% |
|   36–50 years | 17% |
|   51–65 years | 46% |
|   $>65$ years | 35% |
| Farm educational attainment[1] | |
|   High school diploma | 21% |
|   Some college | 30% |
|   Bachelor's degree | 40% |
|   Post-graduate degree | 9% |
| *Data resources/technology use* | |
| High-speed internet access | 80% |
| Use Excel | 43% |
| Use farm data software | 44% |
| Use GPS/autosteer | 92% |
| Use variable rate fertilizer | 71% |
| Use variable rate seeding | 59% |
| Use drone/UAV | 26% |

[1]Highest level of educational attainment among all full-time employees, including owner/operator

respondents are over the age of 50, and 35% are over the age of 65. For comparison, the average age of U.S. farmers is 59 (USDA NASS, 2020). Nearly half of respondents have a bachelor's degree or higher.

In terms of technology use, 80% of the farms surveyed have access to high-speed internet, compared to 74% of rural America as a whole (Wilmoth, 2019). Less than half of respondents reported using either Microsoft Excel or a farm data software product to store, manage, or analyze data. Finally, adoption rates of common precision agriculture technologies are relatively high among survey respondents. GPS guidance/autosteer is used on more than 90% of surveyed farms. 71% of farms surveyed use variable rate fertilizer and 59% use variable rate seeding. Drones/UAVs are used by 26% of the farms surveyed. While these adoption rates are higher than in the general population of U.S. farms (e.g., Lowenberg-DeBoer & Erickson, 2019), they are close to Thompson et al. (2019), who use a similar sampling method focused on commercial sized farms.

## Data collection

Data collection was common with 82% of the 800 respondents colleting yield data, 77% collecting grid or zone soil sample data, and 47% collecting aerial or satellite imagery data. When aggregating across all three data types, 93% of respondents collected at least one of the queried data types.

Comparing demographic variables and data resources/technology use for non-data collectors and data collectors yields several interesting relationships (Table 2). For example, larger farm size is significantly associated with data collection for each of the queried data types (Fig. 2). This relationship is consistent with previous research indicating larger farms are more likely to adopt precision agriculture hardware (e.g., Daberkow & McBride, 2003; Fernandez-Cornejo et al., 2001; Roberts et al., 2004; Schimmelpfennig, 2016; Schimmelpfennig & Lowenberg-DeBoer, 2020). In addition, higher levels of educational attainment are also significantly associated with data collection (Fig. 3), consistent with previous hardware adoption literature (e.g., Fernandez-Cornejo et al., 2001; Roberts et al., 2004).

Data resources and technology use variables were also statistically related with data collection (Table 2). Again, decisions to adopt many of these resources and/or technologies are likely correlated with the decision to collect data. Therefore, it is important to reiterate that the relationships identified in this study are not causal. Moreover, in many cases the direction of the causal pathway cannot be determined as these decisions are made simultaneously. For example, this result does not indicate that farm data software use leads to data collection as it is just as likely that the decision to collect data leads the farm to invest in a data software package. Instead, the associations identified here simply reject the null hypothesis that software use and data collection are independent practices.

Examining the 800 survey responses more closely, just 7% of respondents did not collect yield monitor, soil sample, or satellite/drone imagery data. These non-data collectors identified cost and uncertainty about how to use the data as the two most common impediments to data collection (Fig. 4). Interestingly, privacy concerns (10%) were the least oft cited impediment to data collection among these farms despite well publicized privacy concerns (e.g., American Farm Bureau Federation 2015). When asked about their likelihood to begin collecting data in the future, few non-data collectors indicated a strong propensity to start collecting data in the future suggesting that the move to collecting data is approaching maturity among commercial-scale farms (Fig. 5).

**Table 2** Comparison of data collectors and non-data collectors ($n = 800$)

| Variable | Non-data collectors ($n = 59$) | Data collectors ($n = 741$) | Fisher's exact test[1] |
|---|---|---|---|
| | Percent of respondents | | p-value |
| *Farm/farmer demographics* | | | |
| Farm size | | | 0.000 |
|   1000–1999 acres | 75% | 48% | |
|   2000–4999 acres | 20% | 37% | |
|   ≥000 acres | 5% | 15% | |
| Primary farm owner/operator age | | | 0.741 |
|   <20 years | 0% | <1% | |
|   20–35 years | 0% | 2% | |
|   36–50 years | 14% | 18% | |
|   51–65 years | 46% | 46% | |
|   >65 years | 41% | 34% | |
| Farm educational attainment[2] | | | 0.020 |
|   High school diploma | 37% | 29% | |
|   Some college | 29% | 30% | |
|   Bachelor's degree | 29% | 41% | |
|   Post-graduate degree | 5% | 9% | |
| *Data resources/technology use* | | | |
| High-speed internet access | 69% | 81% | 0.042 |
| Use Excel | 14% | 45% | 0.000 |
| Use farm data software | 5% | 47% | 0.000 |
| Use GPS/autosteer | 73% | 93% | 0.000 |
| Use variable rate fertilizer | 27% | 74% | 0.000 |
| Use variable rate seeding | 22% | 62% | 0.000 |
| Use drone/UAV | 7% | 28% | 0.000 |

[1]Fisher's (1922) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

## Data use

Respondents collecting one or more of the queried data types (yield monitor, grid or zone soil sample, and/or aerial or satellite imagery data), indicated that data they collected commonly influences their seeding rate, nutrient management, and drainage investment decisions (Fig. 6). Nutrient management decisions are the decisions most commonly influenced by farm data, with 93% of data collectors indicating that their data influenced their nutrient management decisions at least somewhat, and more than half indicating that data they collected influenced their nutrient management decisions a lot. Seeding rate decisions are also commonly influenced by farm data, with 81% of those who collected data indicating their data influenced their seeding rate decisions at least somewhat. Of the queried decision types, drainage investment decisions are the
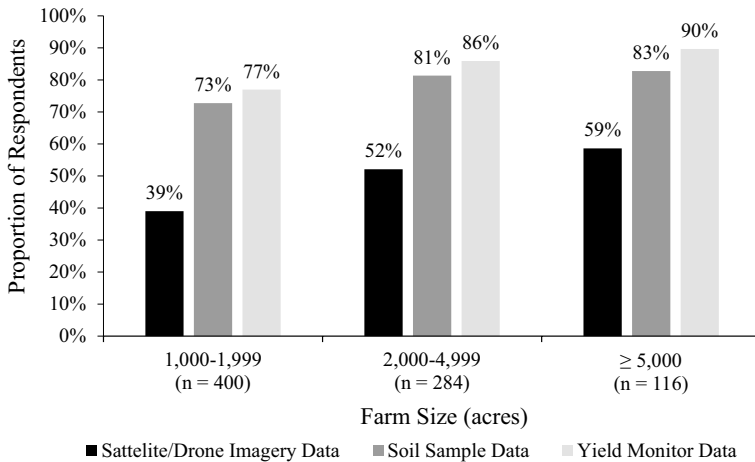
**Fig. 2** Proportion of respondents who collected satellite/drone imagery, soil sample, and yield monitor data by farm size
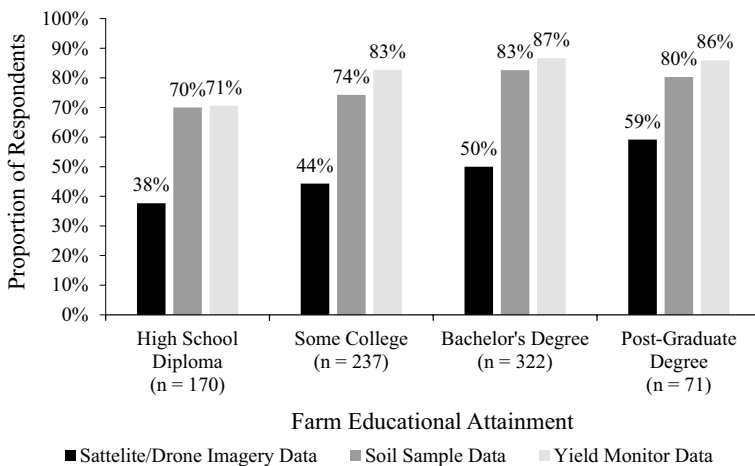


**Fig. 3** Proportion of respondents who collected satellite/drone imagery, soil sample, and yield monitor data by farm education attainment

least likely to be influenced by farm data, although use is still common with 71% of those who collected data indicating it influenced their drainage investment decisions.

Relationships between farm demographics and technology use variables, and the extent to which farm data influences seeding rate, nutrient management, and drainage investment decisions are explored in Tables 3, 4, and 5, respectively. Notice that Tables 3, 4 and 5 include additional variables beyond what is found in Tables 1 and 2. Tables 3, 4 and 5 include variables that are only relevant to the subsample of data collectors. For example, only farms that actually collect data would choose whether or not to designate an individual on their farm to be primarily responsible for collecting,
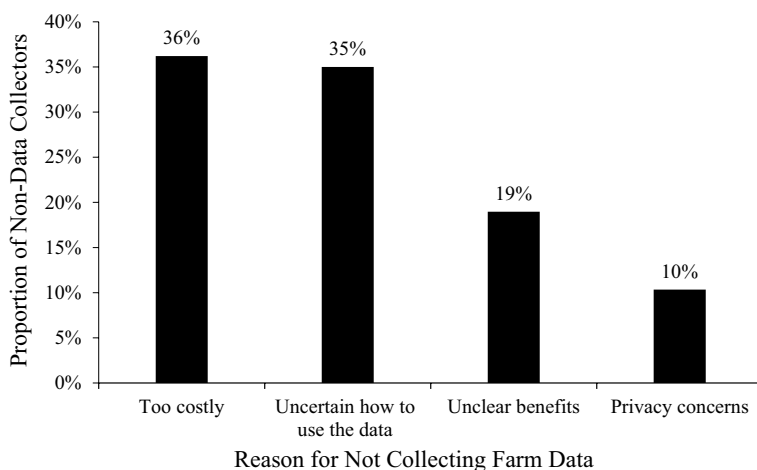
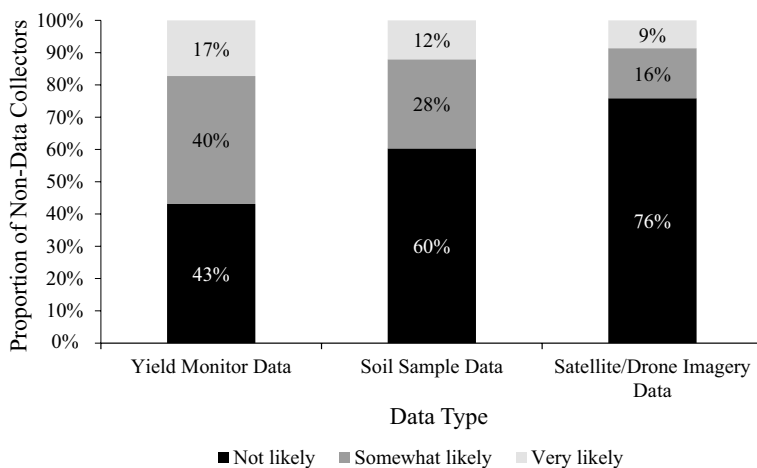**Fig. 4** Reason for not collecting farm data ($n = 59$)



**Fig. 5** Likelihood of non-data collectors to start collecting yield monitor, soil sample, and satellite/drone imagery data in the future ($n = 59$)

managing, and analyzing data. Similarly, only farms that collect data would face the choice to create GPS maps, layer data, or share the data they collect. Hence, while these variables are irrelevant to the data collection decision, they may be associated with the extent to which the farm uses the data it collects or the impact of farm data on farm outcomes.

Demographic variables are associated with the extent to which farm data influences seeding rate and nutrient management decisions. Most notably, larger farm sizes, higher educational attainment, and designation of an employee responsible for collecting, managing and analyzing data are positively related to the level of influence of farm data on seeding rate and nutrient management decisions. However, none of the demographic
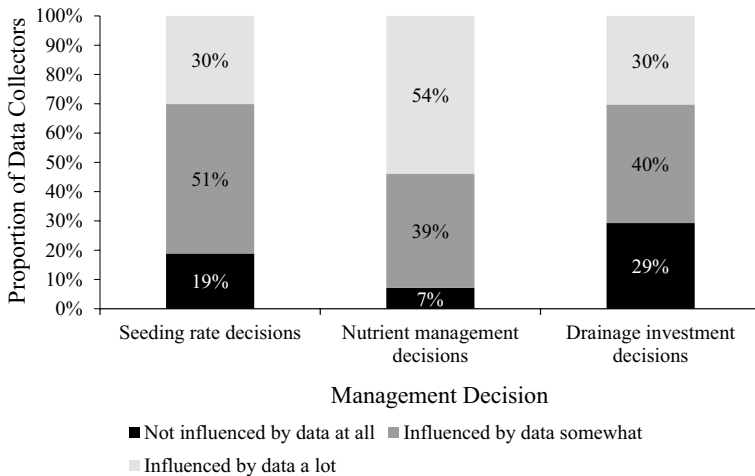
**Fig. 6** Extent to which data collected influenced seeding rate, nutrient management, and drainage investment decisions for data collectors ($n = 741$)

variables were significantly associated with the extent to which farm data influences drainage investment decisions. Nutrient management and seeding rate decisions are different from drainage investment decisions in several ways that may influence this result. Most notably, nutrient management and seeding rate decisions are annual crop management decisions that must be made on all crop acres. Conversely, drainage investment decisions are capital intensive, long-term investment decisions that are typically only evaluated on a small portion of farm acres each year.

Data resources are also associated with the extent to which farm data influences seeding rate, nutrient management, and drainage investment decisions. Farms using Microsoft Excel or farm data software have significantly higher levels of data use. Additionally, creating GPS maps, ability of farm software to layer data and generate recommendations, and following those recommendations all accompany higher levels of data use. This highlights the importance of resources for managing and analyzing farm data. While the landscape of farm data software products is currently vast without a clear-cut market leader, these results suggest that using some form of data software aids farmers in gleaning usable information from their farm's data.

Technology use is also associated with the extent to which data influences farm decisions. VRT fertilizer applications, VRT seed applications, and drone/UAV use are all positively and significantly associated with higher levels of data use. These relationships are unsurprising given that technologies such as VRT require some level of data use to generate VRT prescriptions.

Finally, data sharing is also associated with the extent to which data influences farm decisions. On one hand this is not surprising given the additional value that can be added to farm data by an outside service provider with expertise in managing and analyzing farm data. On the other hand, potential privacy concerns are often an impediment to data sharing (e.g., American Farm Bureau Federation, 2015; Sykuta, 2016; Ferrell, 2017; Miller et al., 2018). While this research does not address these privacy concerns directly, this relationship suggests that when consulted, third-party farm data service providers are adding value to the farm data lifecycle by helping farmers use their data.

**Table 3** Comparison of extent to which farm data influences seeding rate decisions for data collectors ($n = 741$)

| Variable | Extent to which farm data influences seeding rate decisions | | | Fisher's exact test[1] |
|---|---|---|---|---|
| | Not at all ($n = 140$) | Somewhat ($n = 378$) | A lot ($n = 223$) | |
| | Percent of respondents | | | p-value |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.017 |
|   1000–1999 acres | 55% | 48% | 44% | |
|   2000–4999 acres | 38% | 37% | 36% | |
|   ≥ 5000 acres | 7% | 16% | 20% | |
| Primary farm owner/operator age | | | | 0.504 |
|   < 20 years | 1% | <1% | <1% | |
|   20–35 years | 1% | 2% | 2% | |
|   36–50 years | 16% | 16% | 21% | |
|   51–65 years | 41% | 48% | 45% | |
|   > 65 years | 41% | 33% | 32% | |
| Farm educational attainment[2] | | | | 0.041 |
|   High school diploma | 27% | 17% | 20% | |
|   Some college | 33% | 32% | 24% | |
|   Bachelor's degree | 34% | 41% | 46% | |
|   Post-graduate degree | 6% | 10% | 10% | |
| Data employee[3] | 69% | 75% | 83% | 0.003 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 70% | 80% | 88% | 0.000 |
| Use Excel | 36% | 39% | 62% | 0.000 |
| Use farm data software | 30% | 47% | 59% | 0.000 |
|   Farm data software layers data[4] | 38% | 52% | 72% | 0.000 |
|   Farm data software generates recommendations[4] | 50% | 60% | 72% | 0.009 |
|     Follow software recommendations[5] | | | | 0.001 |
|       Do not follow | 9% | 6% | 2% | |
|       Follow somewhat closely | 43% | 64% | 41% | |
|       Follow very closely | 48% | 30% | 57% | |
| Create GPS maps | 52% | 73% | 85% | 0.000 |
| Use GPS/autosteer | 87% | 93% | 96% | 0.008 |
| Use variable rate fertilizer | 62% | 71% | 87% | 0.000 |
| Use variable rate seeding | 33% | 59% | 86% | 0.000 |
| Use drone/UAV | 21% | 27% | 34% | 0.034 |
| *Data sharing* | | | | |
| Share data with service provider | 56% | 77% | 89% | 0.000 |
|   Follow service provider recommendations[6] | | | | 0.004 |
|     Do not follow | 4% | 6% | 7% | |
|     Follow somewhat closely | 67% | 68% | 52% | |
|     Follow very closely | 29% | 26% | 41% | |

**Table 3** (continued)

[1]Fisher's (1922) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[6]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers

## Data impact

Respondents who indicated that farm data influenced their decisions overwhelmingly reported that data-informed decisions increased yield, regardless of the decision (seeding rate, nutrient management, or drainage investment) or the extent to which data influenced the decision (somewhat or a lot) (Fig. 7). It is important to reiterate that yield impacts ascribed to data-informed decision making reported here likely represent respondent perceptions of yield impacts (as opposed to actual yield impacts) given the complexity associated with causally identifying the impact of these data-informed decisions on farm outcomes. Nonetheless, farmer perceptions of these yield impacts are likely to influence future farm decisions (Pannell et al., 2006), and therefore, are an important part of examining the current state of the farm data lifecycle.

It is not particularly surprising that so few respondents indicated that their data-informed decisions decreased yield. On one hand, there is likely survivorship bias associated with this response (Elton et al., 1996). That is, farms that previously perceived yield to decline as a result of data-informed decisions are unlikely to continue using their data to make decisions. There may also be social desirability bias in this distribution of yield impact responses (Fisher, 1993). That is, respondents have the tendency to answer questions in a way that they perceive will be viewed favorably, even if it is not true. Hence, it is possible that some respondents indicated yield increases not because it was reflective of reality, but because they felt that it would be viewed poorly if they indicated that their data-informed decisions did not increase yield. In any case, it is clear that farmers who collect and use farm data perceive their data strategy to be associated with more favorable yield outcomes.

Demographic variables were generally not associated with perceived yield impacts from data-informed seeding rate (Table 6), nutrient management (Table 7), or drainage investment (Table 8) decisions. The one exception was the designation of an employee responsible for collecting, managing, and analyzing data, which was positively and significantly associated with the perception that yield increased as a result of data-informed nutrient management and drainage investment decisions.

Use of data management resources were commonly associated with the perceived impact of data-informed decisions on yield outcomes. Microsoft Excel and farm management software were both associated with perceived yield increases. Similarly, creating GPS maps and the ability of farm software to layer data are both significantly associated with favorable yield impacts.

**Table 4** Comparison of extent to which farm data influences nutrient management decisions for data collectors ($n = 741$)

| Variable | Extent to which farm data influences nutrient management decisions | | | Fisher's exact test[1] |
|---|---|---|---|---|
| | Not at all ($n=53$) | Somewhat ($n=289$) | A lot ($n=399$) | |
| | Percent of respondents | | | p-value |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.072 |
|   1000–1999 acres | 64% | 49% | 45% | |
|   2000–4999 acres | 30% | 36% | 38% | |
|   ≥ 5000 acres | 6% | 15% | 17% | |
| Primary farm owner/operator age | | | | 0.071 |
|   < 20 years | 0% | <1% | <1% | |
|   20–35 years | 6% | 2% | 1% | |
|   36–50 years | 8% | 15% | 21% | |
|   51–65 years | 43% | 48% | 45% | |
|   > 65 years | 43% | 34% | 33% | |
| Farm educational attainment[2] | | | | 0.021 |
|   High school diploma | 28% | 22% | 17% | |
|   Some college | 28% | 33% | 28% | |
|   Bachelor's degree | 42% | 39% | 43% | |
|   Post-graduate degree | 2% | 7% | 12% | |
| Data employee[3] | 57% | 75% | 79% | 0.002 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 79% | 78% | 83% | 0.138 |
| Use Excel | 32% | 37% | 53% | 0.000 |
| Use farm data software | 25% | 42% | 54% | 0.000 |
|   Farm data software layers data[4] | 54% | 50% | 63% | 0.058 |
|   Farm data software generates recommendations[4] | 31% | 58% | 69% | 0.006 |
|     Follow software recommendations[5] | | | | 0.000 |
|       Do not follow | 25% | 8% | 2% | |
|       Follow somewhat closely | 50% | 70% | 44% | |
|       Follow very closely | 25% | 22% | 54% | |
| Create GPS maps | 49% | 62% | 83% | 0.000 |
| Use GPS/autosteer | 87% | 94% | 93% | 0.186 |
| Use variable rate fertilizer | 34% | 66% | 85% | 0.000 |
| Use variable rate seeding | 45% | 56% | 69% | 0.000 |
| Use drone/UAV | 25% | 23% | 32% | 0.046 |
| *Data sharing* | | | | |
| Share data with service provider | 51% | 72% | 84% | 0.000 |
|   Follow service provider recommendations[6] | | | | 0.001 |
|     Do not follow | 11% | 7% | 5% | |
|     Follow somewhat closely | 63% | 72% | 57% | |
|     Follow very closely | 26% | 21% | 38% | |

**Table 4** (continued)

[1]Fisher's ([1922](#)) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[6]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers

Relationships between technology use and perceived impacts of data-informed decisions on yield outcomes were mixed. Farms using VRT fertilizer and VRT seed applications were generally more likely to report favorable yield outcomes. However, it is interesting to note that while VRT seed applications are correlated with yield increases from data informed seeding rate decisions, VRT fertilizer applications are not associated with yield increases from data informed nutrient management decisions.

Finally, sharing data and following the recommendations provided by third-party data service providers is also significantly associated with more favorable perceptions of yield outcomes associated with data-informed decisions. Again, it is not surprising that outside service providers with expertise in managing and analyzing farm data would be able to add value to the farm data lifecycle, especially for annual crop management decisions. Data sharing is not statistically related to perceptions of yield outcomes resulting from data-informed drainage investment decisions.

## Conclusions and implications

The objective of this study is to better understand the current positioning of U.S. commercial-scale corn and soybean farms within the farm data lifecycle. Respondents were explicitly asked about what types of data they collect (i.e., data collection), the degree to which the data they collect influences their farm management decisions (i.e., data use), and finally, how they perceive their data-informed decisions have impacted yield on their farm (i.e., data impact). Results indicate that the majority of commercial-scale U.S. corn and soybean farms collect data, indicate that the data they collect influences their decisions, and perceive yield increases as a result of their data-informed decisions. However, farms vary in intensity of their data usage. When interpreting these results, one must keep in mind that the sample of U.S. commercial corn and soybean farms in this study is intentionally not representative of all U.S. farms, although it is representative of farms producing the majority of these two crops. Further, responses at various levels of the farm data lifecycle may be influenced by selection bias, survivorship bias, and/or social desirability bias. Therefore, results are not generalizable to the broad population of U.S. farms.

Previous research indicated that bundling precision agriculture hardware products provides the best opportunity to maximize the economic returns to a precision agriculture system (Lowenberg-DeBoer, [2003](#); Lambert et al., [2015](#); Schimmelpfennig & Ebel, [2016](#);

**Table 5** Comparison of extent to which farm data influences drainage investment decisions for data collectors (n=741)

| Variable | Extent to which farm data influences drainage investment decisions | | | Fisher's exact test[1] |
|---|---|---|---|---|
| | Not at all (n=217) | Somewhat (n=300) | A lot (n=224) | |
| | Percent of respondents | | | p-value |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.761 |
| 1000–1999 acres | 48% | 50% | 45% | |
| 2000–4999 acres | 36% | 36% | 39% | |
| ≥5000 acres | 16% | 14% | 16% | |
| Primary farm owner/operator age | | | | 0.719 |
| <20 years | 0% | 1% | 0% | |
| 20–35 years | 3% | 2% | 1% | |
| 36–50 years | 18% | 18% | 17% | |
| 51–65 years | 46% | 45% | 48% | |
| >65 years | 34% | 34% | 34% | |
| Farm educational attainment[2] | | | | 0.174 |
| High school diploma | 21% | 21% | 17% | |
| Some college | 31% | 31% | 27% | |
| Bachelor's degree | 38% | 42% | 43% | |
| Post-graduate degree | 10% | 6% | 13% | |
| Data employee[3] | 75% | 76% | 78% | 0.790 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 79% | 81% | 82% | 0.781 |
| Use Excel | 45% | 40% | 53% | 0.014 |
| Use farm data software | 43% | 44% | 55% | 0.022 |
| Farm data software layers data[4] | 44% | 59% | 69% | 0.001 |
| Farm data software generates recommendations[4] | 51% | 71% | 65% | 0.010 |
| Follow software recommendations[5] | | | | 0.415 |
| Do not follow | 8% | 4% | 3% | |
| Follow somewhat closely | 44% | 56% | 51% | |
| Follow very closely | 48% | 39% | 46% | |
| Create GPS maps | 60% | 75% | 81% | 0.000 |
| Use GPS/autosteer | 94% | 93% | 93% | 0.934 |
| Use variable rate fertilizer | 65% | 77% | 79% | 0.003 |
| Use variable rate seeding | 59% | 63% | 65% | 0.380 |
| Use drone/UAV | 24% | 25% | 35% | 0.020 |
| *Data sharing* | | | | |
| Share data with service provider | 72% | 81% | 77% | 0.066 |
| Follow service provider recommendations[6] | | | | 0.331 |
| Do not follow | 8% | 6% | 6% | |
| Follow somewhat closely | 59% | 67% | 59% | |
| Follow very closely | 33% | 27% | 35% | |

**Table 5** (continued)

[1]Fisher's ([1922](#)) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[6]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers
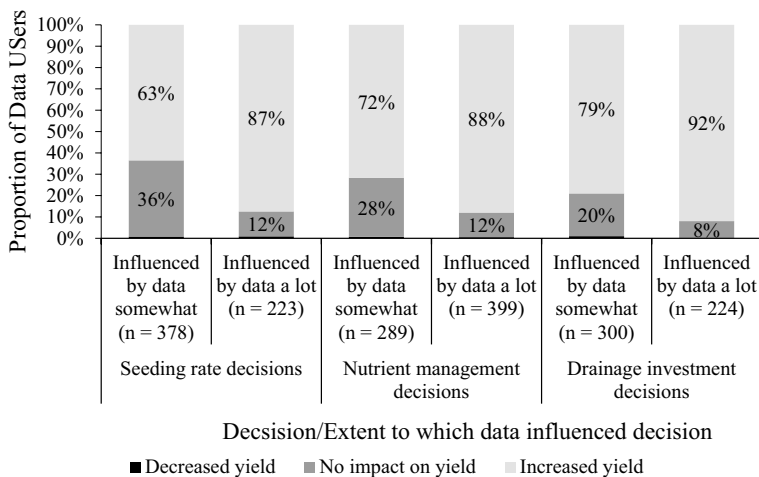


**Fig. 7** Perceived impacts of data-informed seeding rate, nutrient management, and drainage investment decisions on yield outcomes by the extent to which farm data influenced decisions

Miller et al., [2017](#); Thompson et al., [2019](#)). Results of this analysis suggest that it may be important to look beyond hardware when considering the bundle of resources that will maximize the returns to a precision agriculture system. For example, investments in farm software products that can be used to manage and analyze data, including creating GPS maps, layering data sources, and providing recommendations, are associated with progression within the farm data lifecycle. In addition, investments in human capital, either in on-farm employees with designated data responsibilities or in trusted off-farm data services providers that farms are willing to share data with, are also associated with progression within the farm data lifecycle. Therefore, farms wanting to develop and implement a successful farm data strategy should consider the combination of hardware, software, and human capital resources that best fit their farm.

The data revolution in agriculture is fully in motion among U.S. commercial corn and soybean farms as the majority of these farms collect and use data on their farms. Farms that have not yet invested in data management and data analysis resources may be missing out on potential benefits associated with using their farm's data to improve on-farm

**Table 6** Comparison of data-informed seeding rate decisions on perceived yield outcomes for those whose data influenced their seeding rate decisions ($n=601$)

| Variable | Perceived impact of data-informed seeding decisions on yield outcomes | | | |
|---|---|---|---|---|
| | Decreased yield ($n=5$) | No impact on yield ($n=161$) | Increased yield ($n=435$) | Fisher's exact test[1] |
| | Percent of respondents | | | p-value |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.794 |
| 1000–1999 acres | 60% | 49% | 45% | |
| 2000–4999 acres | 20% | 36% | 37% | |
| ≥ 5000 acres | 20% | 15% | 18% | |
| Primary farm owner/operator age | | | | 0.680 |
| < 20 years | 0% | 0% | < 1% | |
| 20–35 years | 0% | 3% | 2% | |
| 36–50 years | 0% | 16% | 19% | |
| 51–65 years | 60% | 50% | 46% | |
| > 65 years | 40% | 31% | 33% | |
| Farm educational attainment[2] | | | | 0.380 |
| High school diploma | 0% | 20% | 18% | |
| Some college | 60% | 24% | 30% | |
| Bachelor's degree | 40% | 47% | 41% | |
| Post-graduate degree | 0% | 8% | 11% | |
| Data employee[3] | 80% | 76% | 79% | 0.798 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 60% | 83% | 84% | 0.274 |
| Use Excel | 40% | 39% | 51% | 0.024 |
| Use farm data software | 40% | 45% | 54% | 0.163 |
| Farm data software layers data[4] | 50% | 49% | 65% | 0.031 |
| Farm data software generates recommendations[4] | 50% | 53% | 69% | 0.023 |
| Follow software recommendations[5] | | | | 0.381 |
| Do not follow | 0% | 8% | 3% | |
| Follow somewhat closely | 0% | 51% | 53% | |
| Follow very closely | 100% | 41% | 44% | |
| Create GPS maps | 60% | 72% | 80% | 0.068 |
| Use GPS/autosteer | 80% | 93% | 95% | 0.177 |
| Use variable rate fertilizer | 60% | 68% | 80% | 0.002 |
| Use variable rate seeding | 80% | 60% | 72% | 0.014 |
| Use drone/UAV | 40% | 31% | 29% | 0.627 |
| *Data sharing* | | | | |
| Share data with service provider | 60% | 72% | 86% | 0.000 |
| Follow service provider recommendations[6] | | | | 0.023 |
| Do not follow | 0% | 13% | 5% | |

**Table 6** (continued)

| Variable | Perceived impact of data-informed seeding decisions on yield outcomes | | | |
| --- | --- | --- | --- | --- |
| | Decreased yield ($n=5$) | No impact on yield ($n=161$) | Increased yield ($n=435$) | Fisher's exact test[1] |
| | Percent of respondents | | | p-value |
| Follow somewhat closely | 33% | 59% | 62% | |
| Follow very closely | 69% | 28% | 33% | |

[1]Fisher's (1922) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[6]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers

decision making. It is important to point out that the benefits are only observable for those that actively manage their data and cannot be assumed as the counterfactual for those that have not invested in data management. Nonetheless, the potential benefits are real. Previous research has shown varying levels of economic benefits associated with precision agriculture use, with small positive economic benefits being the most common conclusion (Griffin et al., 2004; Schimmelpfennig, 2016, 2018; Schimmelpfennig & Ebel, 2016; McFadden, 2017; DeLay et al., 2020; Dhoubhadel, 2021). Even small benefits associated with collection and use of farm data can have big implications for farm structure, as many of the resources discussed here (hardware, software, and human capital) embody fixed costs. Therefore, they are more likely to be adopted on larger farms, where fixed costs can be spread out over more acres. As a result, precision agriculture is hypothesized to have spurred further increases in farm size in recent years (MacDonald et al., 2018). As farms continue to develop and hone their data strategies and product/service providers continue to improve their offerings, digital agriculture could play an important role in continued consolidation within the U.S. farm sector.

**Table 7** Comparison of data-informed nutrient management decisions on perceived yield outcomes for those whose data influenced their nutrient management decisions (n=688)

| Variable | Perceived impact of data-informed nutrient management decisions on yield outcomes | | | Fisher's Exact Test[1] |
|---|---|---|---|---|
| | Decreased yield (n=4) | No impact on yield (n=126) | Increased yield (n=558) | p-value |
| | Percent of respondents | | | |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.931 |
| 1000–1999 acres | 75% | 45% | 47% | |
| 2000–4999 acres | 25% | 38% | 37% | |
| ≥5000 acres | 0% | 17% | 16% | |
| Primary farm owner/operator age | | | | 0.221 |
| <20 years | 0% | 0% | 1% | |
| 20–35 years | 25% | 2% | 1% | |
| 36–50 years | 25% | 17% | 19% | |
| 51–65 years | 25% | 43% | 47% | |
| >65 years | 25% | 39% | 32% | |
| Farm educational attainment[2] | | | | 0.720 |
| High school diploma | 25% | 19% | 19% | |
| Some college | 0% | 30% | 30% | |
| Bachelor's degree | 75% | 44% | 40% | |
| Post-graduate degree | 0% | 7% | 10% | |
| Data employee[3] | 50% | 71% | 79% | 0.060 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 100% | 77% | 82% | 0.381 |
| Use Excel | 25% | 34% | 49% | 0.003 |
| Use farm data software | 0% | 41% | 51% | 0.014 |
| Farm data software layers data[4] | 0% | 44% | 61% | 0.032[5] |
| Farm data software generates recommendations[4] | 0% | 58% | 66% | 0.271[5] |

**Table 7** (continued)

| Variable | Perceived impact of data-informed nutrient management decisions on yield outcomes | | | Fisher's Exact Test[1] |
|---|---|---|---|---|
| | Decreased yield (n = 4) | No impact on yield (n = 126) | Increased yield (n = 558) | p-value |
| | Percent of respondents | | | |
| Follow software recommendations[6] | | | | 0.290[5] |
| Do not follow | 0% | 7% | 4% | |
| Follow somewhat closely | 0% | 60% | 51% | |
| Follow very closely | 0% | 33% | 46% | |
| Create GPS maps | 75% | 71% | 75% | 0.678 |
| Use GPS/autosteer | 100% | 98% | 92% | 0.090 |
| Use variable rate fertilizer | 75% | 71% | 79% | 0.121 |
| Use variable rate seeding | 50% | 55% | 66% | 0.049 |
| Use drone/UAV | 50% | 25% | 28% | 0.392 |
| *Data sharing* | | | | |
| Share data with service provider | 75% | 71% | 81% | 0.063 |
| Follow service provider recommendations[7] | | | | 0.319 |
| Do not follow | 0% | 7% | 6% | |
| Follow somewhat closely | 25% | 69% | 51% | |
| Follow very closely | 50% | 24% | 33% | |

**Table 7** (continued)

[1]Fisher's (1922) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]None of the respondents who perceived data-informed nutrient management decisions to decrease yield reported using a farm software product, and therefore they did not answer whether questions regarding software that layered data, generated recommendations, or the extent to which they followed those recommendations. Therefore, p-values for the Fisher's exact test for these variables represent the null hypothesis of independence between those who indicated data-informed nutrient management decisions did not change yield and increased yield

[6]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[7]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers

**Table 8** Comparison of data-informed drainage investment decisions on perceived yield outcomes for those whose data influenced their drainage investment decisions ($n = 524$)

| Variable | Perceived impact of data-informed drainage investment decisions on yield outcomes | | | |
|---|---|---|---|---|
| | Decreased yield ($n = 3$) | No impact on yield ($n = 78$) | Increased yield ($n = 443$) | Fisher's exact test[1] |
| | Percent of respondents | | | p-value |
| *Farm/farmer demographics* | | | | |
| Farm size | | | | 0.869 |
| 1000–1999 acres | 67% | 51% | 47% | |
| 2000–4999 acres | 33% | 37% | 37% | |
| ≥5000 acres | 0% | 12% | 16% | |
| Primary farm owner/operator age | | | | 0.243 |
| <20 years | 0% | 3% | <1% | |
| 20–35 years | 0% | 0% | 2% | |
| 36–50 years | 33% | 21% | 17% | |
| 51–65 years | 33% | 46% | 46% | |
| >65 years | 33% | 31% | 35% | |
| Farm educational attainment[2] | | | | 0.228 |
| High school diploma | 0% | 24% | 19% | |
| Some college | 33% | 33% | 28% | |
| Bachelor's degree | 33% | 38% | 43% | |
| Post-graduate degree | 33% | 4% | 9% | |
| Data employee[3] | 100% | 67% | 78% | 0.067 |
| *Data resources/technology use* | | | | |
| High-speed internet access | 100% | 79% | 82% | 0.804 |
| Use Excel | 67% | 33% | 48% | 0.026 |
| Use farm data software | 33% | 31% | 52% | 0.001 |
| Farm data software layers data[4] | 100% | 42% | 66% | 0.032 |
| Farm data software generates recommendations[4] | 100% | 67% | 68% | 1.000 |
| Follow software recommendations[5] | | | | 0.175 |
| Do not follow | 0% | 13% | 3% | |
| Follow somewhat closely | 100% | 56% | 54% | |
| Follow very closely | 0% | 31% | 44% | |
| Create GPS maps | 100% | 69% | 79% | 0.093 |
| Use GPS/autosteer | 100% | 90% | 93% | 0.475 |
| Use variable rate fertilizer | 100% | 67% | 79% | 0.035 |
| Use variable rate seeding | 33% | 59% | 65% | 0.291 |
| Use drone/UAV | 67% | 26% | 30% | 0.226 |
| *Data sharing* | | | | |
| Share data with service provider | 100% | 73% | 80% | 0.291 |
| Follow service provider recommendations[6] | | | | 0.046 |
| Do not follow | 33% | 12% | 5% | |

**Table 8** (continued)

| Variable | Perceived impact of data-informed drainage investment decisions on yield outcomes | | | |
|---|---|---|---|---|
| | Decreased yield ($n=3$) | No impact on yield ($n=78$) | Increased yield ($n=443$) | Fisher's exact test[1] |
| | Percent of respondents | | | p-value |
| Follow somewhat closely | 67% | 58% | 64% | |
| Follow very closely | 0% | 30% | 31% | |

[1]Fisher's (1922) exact test tests the null hypothesis that the relative proportions of one categorical variable are independent of a second categorical variable

[2]Highest level of educational attainment among all full-time employees, including owner/operator

[3]Equals one if the farm indicated having an individual on their farm who is primarily responsible for collecting, managing, and analyzing farm data and zero otherwise

[4]Only the subsample of those who used a farm data software were asked about the ability of their software to layer data or generate recommendations

[5]Only the subsample of those who indicated using farm software and that their farm software generated recommendations were asked the extent to which they followed the recommendations produced by their software

[6]Only the subsample of those who indicated sharing data were asked the extent to which they follow recommendations provided by third-party farm data service providers

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Code availability** Data analysis conducted in Stata 16. do files for the current study are available from the corresponding author on reasonable request.

**Declarations**

**Conflict of interest** All authors declare that they have no conflict of interest to declare.

# References

American Farm Bureau Federation. (2015). Privacy and Security Principals for Farm Data. Retrieved 26 March 2020, from https://list.etsi.org/scripts/wa.exe?A3=ind1603&L=AIOTIWG3&E=base64&P= 46619071&B=--_004_7401EDE0FFA3AE4F9F3F3940166B1CB50149B193EMANMB10jdnetde_ &T=application%2Fpdf;%20name=%22PrivacyAndSecurityPrinciplesForFarmData%20(2).pdf% 22&N=PrivacyAndSecurityPrinciplesForFarmData%20(2).pdf&attachment=q&XSS=3

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy, 40,* 79–96.

Daberkow, S. G., & McBride, W. D. (2003). Farm and operator characteristics affecting the awareness and adoption of precision agriculture technologies in the US. *Precision Agriculture, 4,* 163–177. https:// doi.org/10.1023/A:1024557205871

DeLay, N. D., Thompson, N. M., & Mintert, J. R. (2020). Precision Agriculture technology adoption and technical efficiency. Selected paper presented at the 2020 Agricultural and Applied Economics Association Annual Meeting, Kansas City, MO. July 26–28.

Demestichas, K., & Daskalakis, E. (2020). Data lifecycle management in precision agriculture supported by information and communication technology. *Agronomy, 10,* 1648. https://doi.org/10.3390/agronomy10 111648

Dhoubhadel, S. P. (2021). Precision agriculture technologies and farm profitability. *Journal of Agricultural and Resource Economics.* https://doi.org/10.22004/ag.econ.303598

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and mutual fund performance. *The Review of Financial Studies, 9,* 1097–1120.

Ferrell, S. L. (2017). Technology in farming: Data driven agriculture. Written testimony before United States Senate Committee on Commerce, Science, and Transportation Subcommittee on Consumer Protection, Product Safety, Insurance, and Data Security. Retrieved from https://www.commerce.senate. gov/services/files/D85C0F87-2358-4452-9216-882B61DE3856

Fernandez-Cornejo, J., Daberkow, S. G., & McBride, W. D. (2001). Decomposing the size effect on the adoption of innovations: Agrobiotechnology and precision agriculture. *AgBio Forum, 4,* 124–136.

Fisher, R. A. (1922). On the interpretation of $\chi^2$ form contingency tables, and the calculation of P. *Journal of the Royal Statistical Society, 85,* 87–94.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20,* 303–315.

Griffin, T. W., Lowenberg-DeBoer, J., Lambert, D. M., Peone, J., Payne, T., & Daberkow, S. G. (2004). Adoption, profitability, and making better use of precision farming data. Staff Paper #04–06, Department of Agricultural Economics, Purdue University. https://doi.org/10.22004/ag.econ.28615

Griffin, T. W. N. J., Miller, J., Bergtold, A., Shanoyan, A., Sharda, & Ciampitti, I. A. (2017). Farm's sequence of adoption of information-intensive precision agriculture technology. *Applied Engineering in Agriculture, 33,* 521–527. https://doi.org/10.13031/aea.12228

Khanna, M., Epouhe, O. F., & Hornbaker, R. (1999). Site-specific crop management: Adoption patterns and incentives. *Review of Agricultural Economics, 21*, 455–472. https://doi.org/10.2307/1349891

Lambert, D. M., Paudel, K. P., & Larson, J. A. (2015). Bundled Adoption of precision agriculture technologies by cotton producers. *Journal of Agricultural and Resource Economics, 40*, 325–345.

Lowenberg-DeBoer, J. (2003). Precision farming or convenience agriculture. Paper presented at the 11th Australian agronomy conference, Geelong, Victoria, Australia, February 2–6.

Lowenberg-DeBoer, J., & Erickson, B. (2019). Setting the record straight on precision agriculture adoption. *Agronomy Journal, 111,* 15152–11569. https://doi.org/10.2134/agronj2018.12.0779

MacDonald, J. M., Hoppe, R. A., & Newton, D. (2018). Three decades of consolidation in U.S. agriculture. Economic Bulletin Number 189, Economic Research Service. Retrieved from https://www.ers.usda. gov/webdocs/publications/88057/eib-189.pdf

McFadden, J. R. (2017). Yield maps, soil maps, and technical efficiency: Evidence from U.S. corn fields. Paper presented at the agricultural and applied economics association annual meeting, Chicago, IL. July 30-Austust 1.

Miller, N. J., Griffin, T. W., Bergtold, J., Ciampitti, I. A., & Sharda, A. (2017). Farmers' adoption path of precision agriculture technology. *Advances in Animal Bioscience, 8*, 708–712. https://doi.org/10.1017/ S2040470017000528

Miller, N. J., Griffin, T. W., Ciampitti, I. A., & Sharda, A. (2019). Farm adoption of embodied knowledge and information intensive precision agriculture technology bundles. *Precision Agriculture, 20*, 348–361. https://doi.org/10.1007/s11119-018-9611-4

Miller, N. J., Griffin, T. W., Goeringer, P., Ellixson, A., & Shanoyan, A. (2018). Estimating value, damages, and remedies when farm data are misappropriated. *Choices, 33*, 1–8.

Pannell, D. J., Marshall, G. R., Barr, N., Curtis, A., Vanclay, F., & Wilkinson, R. (2006). Understanding and promoting adoption of conservation practices by rural landholders. *Australian Journal of Experimental Agriculture, 46*, 1407–1424. https://doi.org/10.1071/EA05037

Plume, K. (2014). High-tech, U.S. farm machines harvest big data, reap privacy worries. Reuters. Retrieved from https://www.reuters.com/article/usa-farming-data-idUSL2N0N11U720140409

Pope, M., & Sonka, S. (2020). Quantifying the economic benefits of on-farm digital technologies. *farmdoc daily* 10:40. Retrieved from https://farmdocdaily.illinois.edu/2020/03/quantifying-the-economic-benefits-of-on-farm-digital-technologies.html

Roberts, R. K., English, B. C., Larson, J. A., Cochran, R. L., Goodman, W. R., Larkin, S. L., Marra, M. C., Martin, S. W., Shurley, W. D., & Reeves, J. M. (2004). Adoption of site-specific information and variable-rate technologies in cotton precision farming. *Journal of Agricultural and Applied Economics, 36*, 143–158. https://doi.org/10.22004/ag.econ.42943

Schimmelpfennig, D. (2016). Farm profits and adoption of precision agriculture. U.S. Department of Agriculture, Economic Research Service, Report Number 217. Retrieved from https://www.ers.usda.gov/webdocs/publications/80326/err-217.pdf?v=0

Schimmelpfennig, D. (2018). Crop production costs, profits, and ecosystem stewardship with precision agriculture. *Journal of Agricultural and Applied Economics, 50*, 81–103. https://doi.org/10.1017/aae.2017.23

Schimmelpfennig, D., & Lowenberg-DeBoer, J. (2020). Farm types and precision agriculture adoption: crops, regions, soil variability, and farm size. Global Institute for Agri-Tech Economics Working Paper 01–20.

Schimmelpfennig, D., & Ebel, R. (2016). Sequential adoption and cost savings from precision agriculture. *Journal of Agricultural and Resource Economics, 41*, 97–115.

Skyuta, M. E. (2016). Big data in agriculture: Property rights, privacy and competition in Ag data services. *International Food and Agribusiness Management Review, 19*, 57–74.

Thompson, N. M., Bir, C., Widmar, D. A., & Mintert, J. R. (2019). Farmer perceptions of precision agriculture technology benefits. *Journal of Agricultural and Applied Economics, 51*, 1–22.

USDA Economic Research Service. (2020). USDA agricultural resource management survey farm financial and crop production practices tailored reports. https://data.ers.usda.gov/reports.aspx?ID=17883.

USDA National Agricultural Statistic Service. (2020). "Quick Stats." https://quickstats.nass.usda.gov/

Weersink, A., Fraser, E., Pannell, D., Duncan, E., & Rotz, S. (2018). Opportunities and challenges for big data in agricultural and environmental analysis. *Annual Review of Resource Economics, 10*, 19–37. https://doi.org/10.1146/annurev-resource-100516-053654

Wilmoth, D. (2019). Accessing the internet in rural America. Issue Brief Number 15, U.S. Small Business Administration, Office of Advocacy. Retrieved from https://cdn.advocacy.sba.gov/wp-content/uploads/2019/11/07105617/Rural-Internet-Accesss.pdf

Woodard, J. D., Sherrick, B. J., Atwood, D. M., Blair, R., Fogel, G., Goeser, N., et al. (2018). The power of agricultural data. *Science, 362,* 410–411. https://doi.org/10.1126/science.aav5002

Zhou, X. V., English, B. C., Larson, J. A., Lambert, D. M., Roberts, R. K., Boyer, C. N., et al. (2017). Precision farming adoption trends in the southern U.S. *Journal of Cotton Science, 21,* 143–155.