The International Society of Precision Agriculture presents the
15th International Conference on
Precision Agriculture
26–29 JUNE 2022
Minneapolis Marriott City Center | Minneapolis, Minnesota USA

Use of precision technologies to conduct successful within-field, on-farm trials

Andrew J. Krmenec[1,2], Mark Stelford[1]

[1] Premier Crop Systems, Iowa, USA
[2] Northern Illinois University, Illinois, USA

A paper from the Proceedings of the
15th International Conference on Precision Agriculture
June 26-29, 2022
Minneapolis, Minnesota, United States

Abstract.

Performing randomized replicated trials in row crop field environments has the potential to increase crop production in environmentally sustainable ways. Successful implementation requires an understanding of implement capabilities and sources of potential systematic error, including operator error. We examine the types of errors encountered from six years of managing OFEs in production. Specific OFE errors are reduced to seven general classes: Cancelled Experiment, Crop Failure/Loss, V.R. Application Problems, Execution/Completion, Controller/Logger Problems, Protocol Failure, and Geography Error. Within each class there are errors that are fatal to the experiment, invalidating its use for science, and errors that are potentially resolvable. We use these examples to illustrate how OFEs might fail the criteria for scientific knowledge building and to suggest the need for metadata standards that would give agricultural science confidence in using OFE data and findings.

Keywords.

On-farm precision experiments, Experiment quality, Knowledge transferability, Metadata

---

Introduction

On-farm cropping experimentation has traditionally been largely informal, in the sense that it has not been designed for rigorous scientific analysis (Lockeretz 1995). Precision technologies now give farmers and agronomists the ability to plan and implement in-field experiments based on formal agronomic science protocols, such as the randomized complete block design, Latin square, and strip trial. With proper planning and execution, these experiments can be conducted on smaller footprints and at greater volume within a farmer's operations.

The on-farm experimentation literature widely recognizes the value of these experiments for farmer/grower learning and the potential value as well for knowledge building in agronomic science (Langford 2022). The differences between farmer learning and knowledge building for science are not trivial. A farmer may accept simple average yield by treatment statistics as valued outcome from an experiment, for example, whereas science expects a rigorous test. Science is highly conservative in its criterion for rejecting no outcome difference between treatments, whereas the farmer may be willing to accept a higher degree of risk. Thompson et al. (2019) suggest that such differences are not a burden to on-farm experimentation however, as farmers also value learning about the research process.

Cho et al. (2021), Lawes and Bramley (2012), and Mallarino et al. (2011) present methods for implementation of OFE through strip trials. Bullock et al. (2020), Bullock et al. (2019) and Trevisan et al. (2021) demonstrate the use of Latin square experiment designs. Both of these experiment designs are spatially extensive, especially the whole-field Latin square, and have significant economic implications for the farmer in terms of foregone yield and the costs of suboptimal or ineffective treatments. Randomized complete block designs have the potential to reduce OFE footprints and experiment costs but require a level of uniformity of background environment conditions not easily controlled or verified (Sumner 2015).

With precision technologies capable of fine-scale implement control and the proven adaptability of experiment designs to on-farm use, it would seem that a well-structured OFE approach could accelerate knowledge building in precision agriculture and agronomic science (Langford 2022; Willers et al. 2008; Bramley and Grains Research and Development Corporation 1999). Tanaka (2021) warns, however, that data quality questions remain. This article considers the range of factors affecting data quality in on-farm experiments. We draw upon six years of experience analyzing outcomes from randomized block experiments developed primarily for plant population and nutrient rate treatments. The lessons learned regarding data quality are applicable to any OFE design and highlight the importance of careful metadata documentation.

Error, Systematic Error and Experiment Quality

Experiment quality fundamentally boils down to control, precision and error. Control captures the degree to which the experimenter can regulate the environment and context in which the experiment takes place. It includes having the ability to fix treatment levels–including a "control" treatment, calibrate equipment and measurement instruments, and limit the effect of outside/non-treatment influences. Precision can be thought of the sensitivity of instruments or degree of tuning available to experiment processes. A more finely tuned instrument can discriminate smaller increments of meaningful difference in an outcome, treatment level, or process parameter. The technologies of precision agriculture have demonstrated a high degree of capability, both in terms of controllability and precision.

Error is the measure of inaccuracy. It can be systematic or asystematic. Asystematic error has no recognizable pattern or relation to any of the treatments/factors or processes involved in the experiment and is typically referred to as random error. Its source is not identifiable and as a consequence it is not correctable. Systematic error, also known as bias, has some pattern. It is diagnosable if we have access to a truth state. Systematic error may be consistent and thus addressable, or inconsistent. For example, a given planter set up for corn may not be able to reach and hold a target treatment rate of 118,000 seeds per hectare (~ 48000 seeds/acre), instead maxing out and holding steady around 112,000 seeds per hectare (~ 45000 seeds/acre). Rather than invalidate the experiment, this error is addressable post-experiment by adjusting the target treatment to 112,000. If the planter maxes out at 112,000 seeds/ha., but it's recorded as-planted rate fluctuates widely and randomly in the vicinity of 112,000, the error is systematic, inconsistent and unaddressable.

Errors of any form reduce the information content of experiment data and the reliability of calculated statistics. The worst case is error that is systematic but inconsistent. In that instance statistics derived from the experiment are biased and untrustworthy. In all cases, larger error results in statistics with larger sampling variances and a greater likelihood of failure to recognize true treatment effects. OFE data, like any experiment data, contain some degree of error. The source(s) and type(s) of error present ultimately affect the value of those data for knowledge generation.

Study Data and OFE Design

The analysis to follow examines the nature of errors identified in quality control of randomized block OFEs conducted between 2016 and 2021. The experiments were created in a tool available to consulting agronomists either employed by our company or employed by client firms through the upper Midwest U.S.A. Farmers engaged in these OFEs typically plant a soybeans-corn rotation or a soybeans-corn-wheat rotation. Table 1 shows the volume of experiments created by year with the vast majority being either plant population or nutrient rate experiments. These records demonstrate strong farmer interest in conducting OFEs.

Table 1. Record of OFE experience

Crop year	2016	2017	2018	2019	2020	2021
# OFE planned	338	901	1971	2998	4606	4969
annual growth rate		267%	219%	152%	154%	108%
% "error-free"	50.2%	55.7%	55.9%	54.1%	45.3%	61.6%

The protocol for OFEs created through our tool calls for only one treatment factor, applied by the same implement, and harvested by the same harvester. Treatment applications must occur within a 24-hour period, and harvest of the entire OFE likewise must be conducted within a 24-hour period. Each treatment level receives at least five replicates in the experiment footprint, with replicate width and length determined by the type of experiment, swath dimensions of the applicator and harvester, and orientation of harvest relative to planting. Dry nutrient applications have longer replicate run lengths due to the lesser precision those technologies have for adjusting application rates. Small grain harvesting in the study region are often oriented at a slight (3-5 degree) offset to planted rows. OFE replicates in soybeans and wheat are therefore generally wider than a multiple of the least common denominator of applicator and harvester swath widths. The typical OFE footprint for a plant population study with 5 randomly situated replicates at each of three treatment rates is roughly 1 hectare (2.5 acres). Farmers may increase the number of replicates or the number of treatment levels beyond these minima.

There are several critical requirements in our OFE protocol that drive diagnoses of errors and experiment quality. First, we require that the entire experiment footprint lie within a uniform crop/soil environment; it cannot overlap into headlands or untilled waterways. Second, we require that all other management practices within the footprint of the OFE be uniform – same equipment, same product, same rate, same timing. Third, applicator and harvester passes need to be conducted at "production speed" and in sequence with areas adjacent to the experiment. Fourth, applicator and harvester swaths need to be full throughout the length of the replicate, and not intersect any other applicator or harvester swath. Table 1 shows that a good proportion of experiments planned and conducted through our system achieved "error-free" status.

A Typology of Observed OFE Errors

A complete quality control of OFE recognizes that errors affecting experiment quality can occur anywhere between creation of the experiment and analysis of final harvest data. Since execution of the experiment is conducted by farm staff, and not the researcher, each step in the OFE process needs to be carefully examined. Are the treatment levels for the experiment appropriate to that type of experiment and that crop? Are treatment prescriptions properly converted to units the applicator's control system understands? Does the applicator produce the desired target treatment rates? Have all experiment protocols been followed? Is the experiment complete through to recorded yield data? Are the yield measurements reliable? Have they been properly calibrated?

The quality control process we follow with OFEs created and analyzed in our system consists of three separate evaluations. First, as-applied data are evaluated with respect to consistency of the as-applied rates (products or timings), correspondence to target and ℞ rates, date, and application path geography. Similar analyses are done separately for harvest data, with a focus on data extremes, extreme point-to-point variability, GPS drift, and path and swath overlaps. A third quality control evaluation is done upon intersection of the as-applied and harvest data. Here we identify and eliminate transition zones between replicates with dissimilar treatments, check whether as-applied and harvest orientations are consistent, and consistent with the OFE plan, and assess the intersections of as-applied and harvest data for geography match. A summary of error types identified in these quality control processes is presented in Table 2.

Two classes of OFE error result in invalid experiments. The first occurs either when a planned experiment is cancelled (not conducted) and the second when there is crop failure or loss. In either case, there is no valid experiment data. Most of the other classes of OFE error have a mix of error types either fatal to the experiment or potentially resolvable, post-experiment. Failure to follow experiment protocol on using the same equipment, for example, are fatal because there is no way to calibrate one implement against another – be it in terms of as-applied or harvest data values. Other fatal, OFE-invalidating errors include: harvest monitor failure, resulting in unreliable or missing harvest data; missing prescription data, preventing full assessment of as-applied data quality; crop changes; multiple treatments within experiment footprint, as opposed to just a single factor; application or harvest on different days; and cases where the treatment is applied or the crop harvested in dissimilar directions.

Other types of errors in Table 2 may be fully or partially resolvable. If a small portion of harvest data is missing, for example, one solution to preserving some knowledge gain from the experiment might be to delete the treatment replicate(s) and randomly delete replicates from other treatment levels so as to preserve a balanced design. In the case where treatment levels are too similar (insufficient rate variability), a solution might be to eliminate the middle rate(s) and only evaluate experiment data from replicates of the highest and lowest rates. This method of partial resolution may be preferred also when as-applied rates are inconsistent (high C.V.) or the applicator fails to hold a treatment rate. Whether resolved by this means or any other, the presence of either of these latter errors should lessen our confidence in analyses derived from those experiments.

Most of the remaining errors in Table 2 are geography errors. Geography errors present as mismatches between the planned footprint of the OFE and its execution. Examples include overlapping headlands or a waterway, overlap of applicator paths or harvester paths, and OFEs overlapping. Poorly calibrated machinery look-ahead distances, excessive GPS drift, and the use of machinery with larger or smaller swath widths than those specified in OFE planning also produce geography errors. Guidance lines are a natural solution to preventing several of these errors as long as they are followed. On large operations, where short-term contract operators are hired to help with planting and harvest, time urgency or operator discretion may prevail over OFE trafficking plans based on guidance lines.

Table 2. Classification and Severity of Major OFE Errors

OFE Error Class	*Description (example)*	*Severity*
Cancelled Experiment	Experiment not conducted	fatal
Crop Failure/Loss	Drown-out	fatal
	Wind damage (derecho loss, stand lodging)	fatal
	Replant	fatal
V.R. Application Problems	Insufficient rate variability (treatment levels too close)	potentially resolvable
	Flat rate, uniform treatment	fatal
	Incorrect rates/products/timing	potentially resolvable
	Applicator didn't hold rate	potentially resolvable
	High as-applied C.V.	potentially resolvable
Execution/Completion	As-applied data incomplete	potentially resolvable
	Partial application execution	potentially resolvable
	Missing ℞	fatal
	Harvest data incomplete	potentially resolvable
Controller/Logger Problems	Look-ahead error	potentially resolvable
	Harvest monitor error	fatal
Protocol Failure	Crop change (soybeans instead of corn)	fatal
	Multiple varieties or products	fatal
	Wrong application/harvest direction (vs. OFE plan)	potentially resolvable
	Wrong machinery dimension (vs. OFE plan)	potentially resolvable
	Extreme application angle	potentially resolvable
	Extreme harvest angle	potentially resolvable
	Application not on same day	fatal
	Harvest not on same day	fatal
	Multiple application machines	fatal
	Multiple harvesters	fatal
Geography Error	Application or harvest path offset	potentially resolvable
	Experiments overlap	potentially resolvable
	Application path overlap	potentially resolvable
	Harvest path overlap	potentially resolvable
	Headland overlap	potentially resolvable
	Waterway overlap	potentially resolvable
	Multiple application/harvest orientations	fatal

Figures 1a-d illustrate the consequences of several types of geography errors. Experiment replicates are shaded according to their (randomly) assigned treatment level. The point data are locations of as-applied recordings, again shaded according to their assigned target rates. Figures 1a and b also include series of black as-applied points. These are invalid; in (a) because the implement operator ignored experiment guidance lines and in (b) because the applicator used was ~3m (10 ft.) narrower than what was specified in experiment creation. Fortunately, there was enough overlap between harvest points and valid as-applied points in both examples to complete experiment analysis – but not without loss of data.

Figure 1c shows the consequence of an improperly set look-ahead distance on the applicator's controller. Notice that as-applied treatment levels are offset west of replicate boundaries in the northern-most applicator swath. The same westward offset occurs in swaths 3, 5, 7, 9, whereas treatment levels are offset eastward in swaths 2, 4, 6, 8. This type of error is resolvable by increasing the width of the transition zones between unlike replicates; however, there is a cost in terms of reduced replicate size (run length). Figure 1d shows another form of geography error, in this case due to the operator stopping field operations over the footprint of the experiment and returning to singularly apply the treatment to the OFE later. We classify this as a protocol error as well as a geography error. The whole experiment is therefore invalid. Contrast figures 1a-d with figure 2, which shows a near ideally executed experiment.

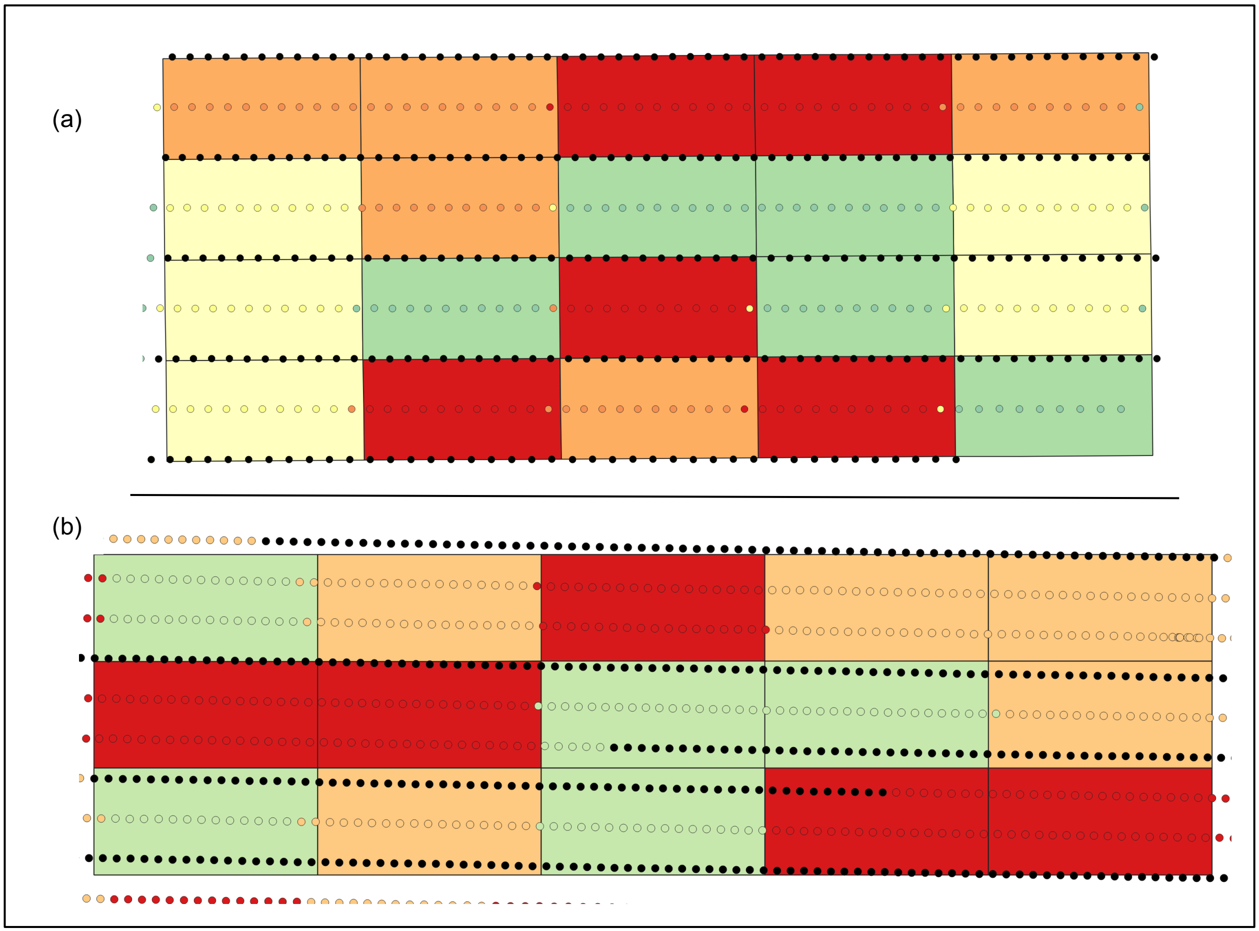

Fig. 1 Geography Errors: (a) application path did not follow guidance lines; (b) applicator swath width greater than planned swath width

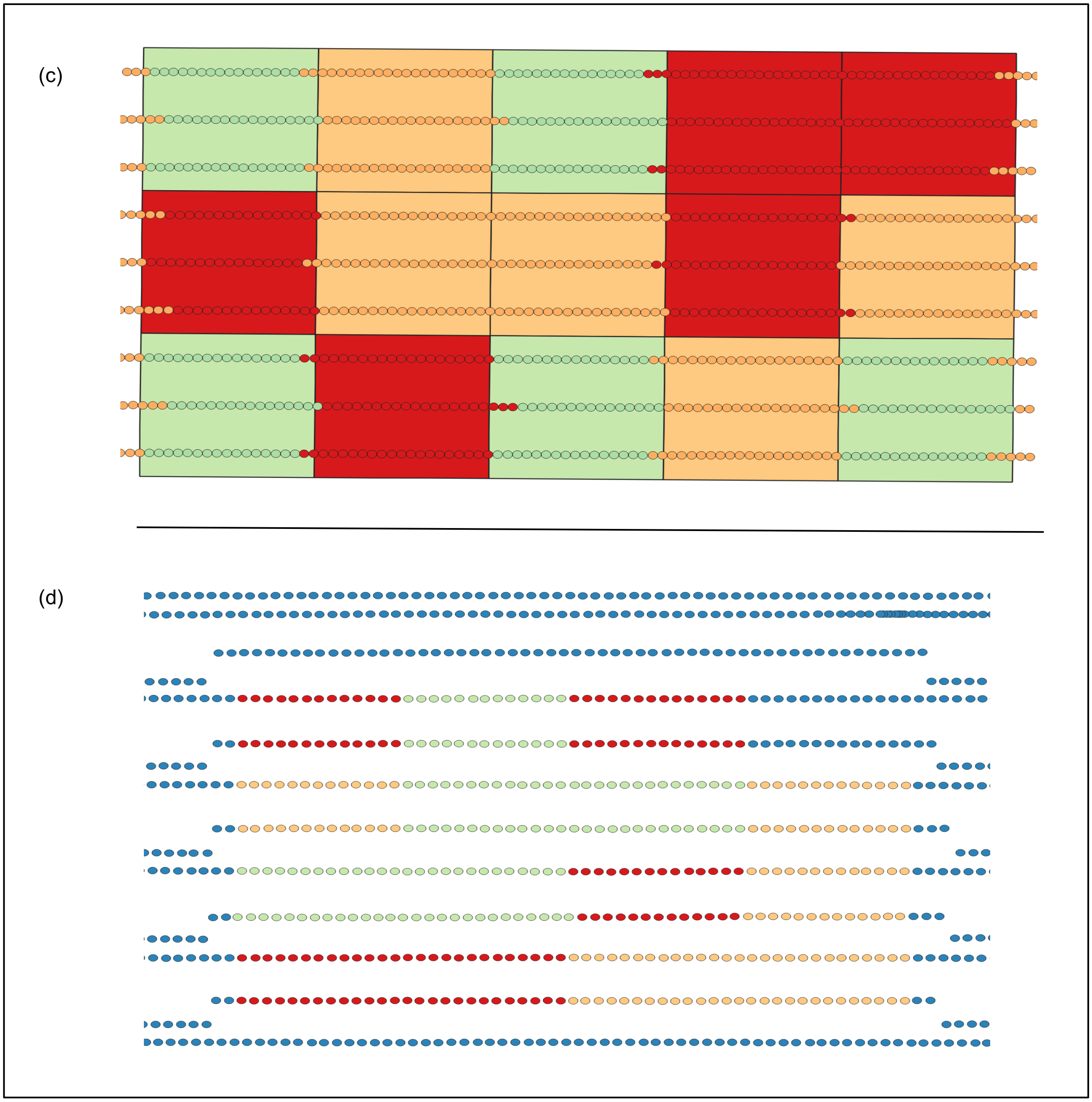

Fig. 1 cont'd. Geography Errors: (c) incorrect look-ahead distance; (d) application conducted separate from field operations

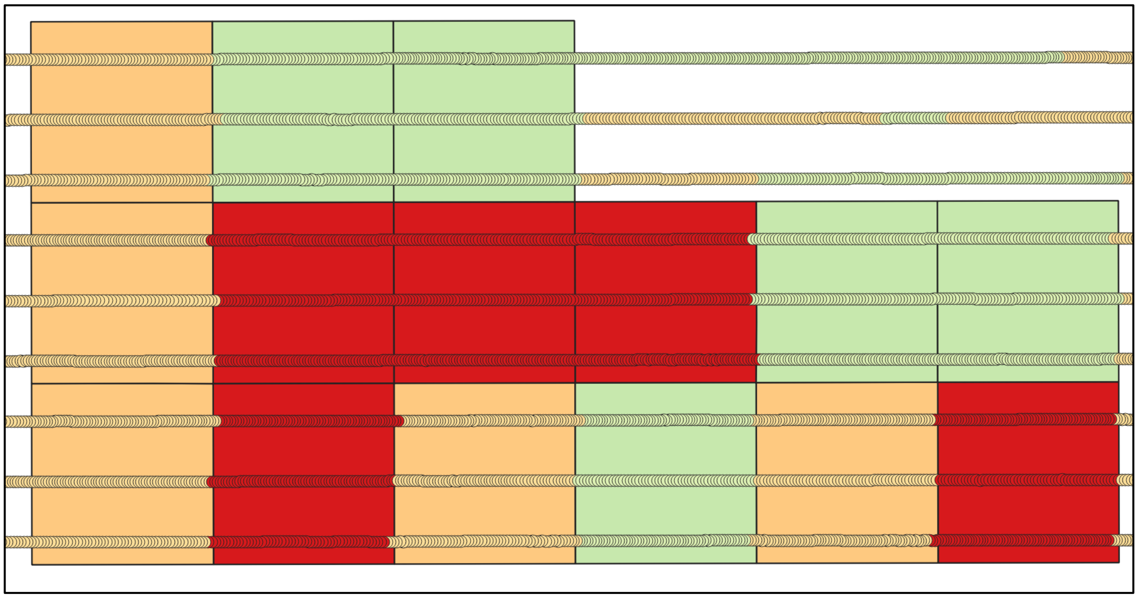

Fig. 2 A near-ideal applicator execution

Historically, the proportion of OFEs conducted through our system that were identified as error-free ranges from 45% to 61% (Table 1). As we quality control as-applied data, harvest data and the spatial intersection of as-applied and harvest, a given experiment can be diagnosed with multiple errors. Figure 3 shows the relative distribution of OFE errors, by error class and year. In the first years of running these experiments, V.R. application problems dominated. The share of those error types has diminished over time. Note as well that controller/logger problems have essentially disappeared. As farmers have grown more comfortable with setting up and executing OFEs, both V.R. application and controller/logger problems should be expected to decline. Crop loss errors are primarily weather related and thus are variable from year to year.

The largest class of OFE errors over recent years has been protocol errors. Between 2017 and 2020 roughly 40-45% of all diagnosed errors were classed as protocol failure. In 2021, that proportion jumped to 65%. The single most dominant cause of protocol failure over all years has been multiple varieties or multiple products (Table 3). Some of the cases of multiple product violations of protocol are associated with headlands overlap, as farmers in some areas plant a different variety in more heavily trafficked headlands. The most recent year of our data (2021) showed a jump in protocol failure resulting from experiment implementation with machinery different in swath width from what was specified in the experiment creation process (see Figure 1b). Much of the increased protocol failures due to these two specific errors is likely due to phenomenal growth in the number of new farmers integrating OFEs into their annual cropping plans and the learning curve associated with our experiment protocol.

Table 3. Protocol failure and geography error by year, as % of all errors, and main types of protocol errors as % of all protocol failure.

	2016	2017	2018	2019	2020	2021
Protocol Failure	26%	40%	47%	41%	45%	65%
Multiple varieties/products	89%	91%	91%	66%	80%	69%
Wrong machinery dimensions	3%	2%	1%	9%	3%	23%
Geography Error	20%	12%	6%	9%	15%	5%

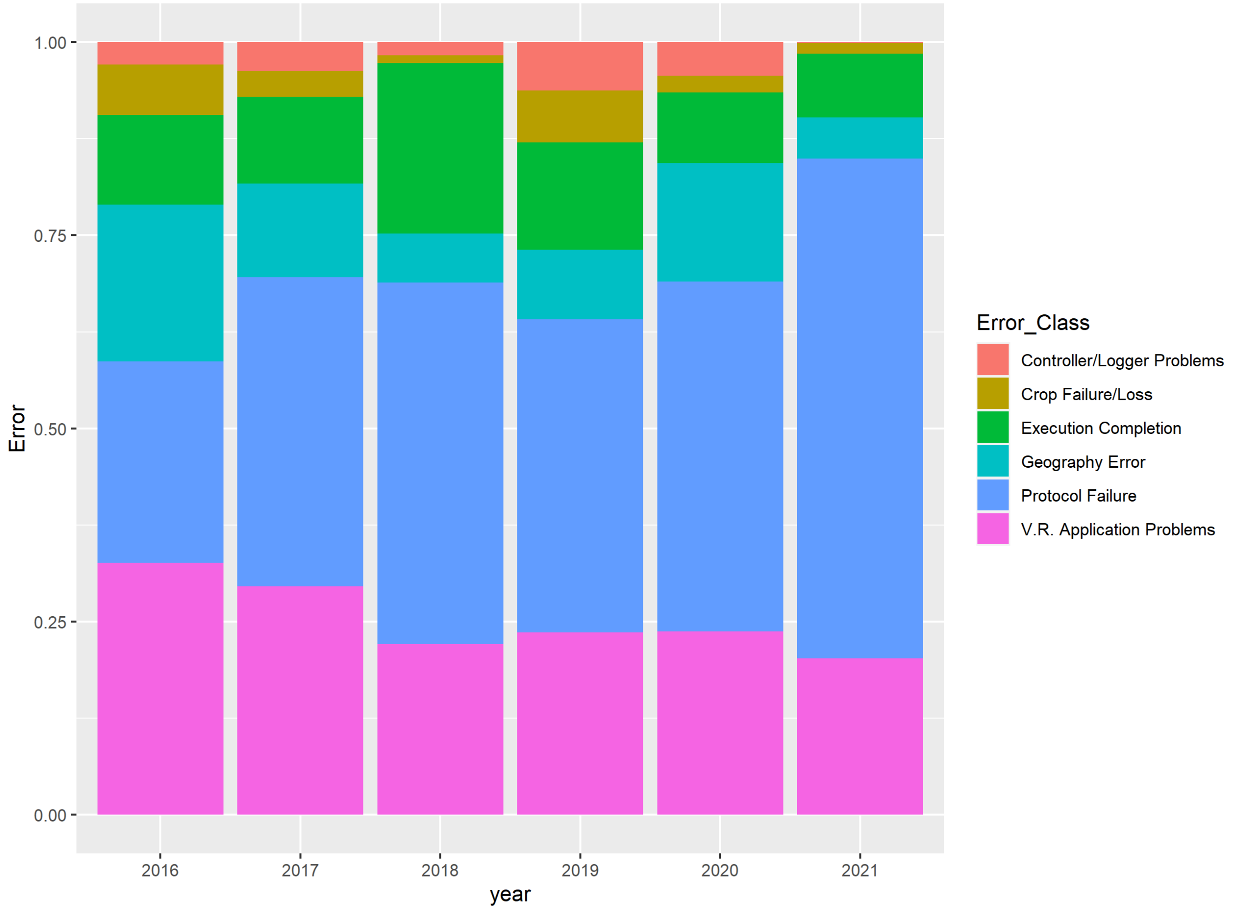

Fig. 3 Relative distribution of OFE error classes, by year

Discussion

On-farm experiments offer more than just opportunity for farmer learning. Baars (2011), Reetz (2011) and others have advocated for researcher-led OFEs as a means to realize more value from these experiments. Indeed, Carton et al. (2022) and Thompson et al. (2019) have shown that researcher-farmer collaborations increase farmers' comfort with scientific research and can lead to the ability to conduct even more sophisticated OFEs. Kraaijvanger and Veldkamp (2015) go further and argue that researchers need the experiential local knowledge of farmers to ensure that they have a complete understanding of the experiment environment and context.

Science has well-established criteria for knowledge generation that transcend individual researchers and indeed whole domains of knowledge inquiry. Of particular importance at the experiment level are accuracy and precision. Hansson (2019) notes that the advantages of farmer produced experiments may be offset by precision and reliability issues, while Steinke et al. (2017) raise questions about the accuracy of OFE data. These are justifiably valid concerns as our own analysis has shown. If agricultural science is to reap true value out of OFEs, whether researcher led or farmer led, it needs the ability to fully assess the experiment design, protocol, and data against common standards (cf., Alesso et al. 2019).

Precision agriculture and the OFE community have steered clear of advocating any particular design for on-farm experiments as we well know that an experiment is designed to answer its research question. That doesn't mean, however, that these communities should refrain from documenting the metadata information they need in order to assess OFE quality. Rather, it is to the benefit of science, farmers and society that standards be set and well publicized. An experiment's metadata should begin with the experiment question. It should include detailed information on the experiment design and underlying implementation protocol. The metadata should all document all data measurement methods, including calibration procedures. If data are quality controlled, those processes should also be documented along with any parameters or standards used in assessing quality. To address reliability concerns, C.V. information for as-applied and harvest data need to be included in experiment metadata – ideally by treatment and treatment replicate. Care must be taken to respect and maintain the farmer's ownership of the raw experiment data while conforming to FAIR principles (see, Peng et al. 2021). Similar to geospatial metadata standards (Danko 2012), we suggest that the metadata identify what data aggregation measures can be shared, to whom, and under what end-use conditions.

Conclusion

To increase the value of OFE towards scientific knowledge generation, scientists need to have the ability to fully review experiment design, implementation protocol, data measurement and quality control, and both as-applied and harvest reliability. This paper presented a summary of diagnosed experiment errors from six years of creating and managing randomized block OFEs. Certain errors invalidate the use of the experiment for both farmer learning and science. Other errors are potentially resolvable, thus returning value to the farmer. But adjustments for those errors carry reliability and confidence implications. As yet there are no metadata standards for documenting OFEs and data quality assessment operations. This limits OFE utility to science. The precision agriculture and OFE communities could help accelerate the acceptance and use of on-farm experiments by developing experiment metadata standards.

References

Alesso, C. A., Cipriotti, P. A., Bollero, G. A., & Martin, N. F. (2019). Experimental Designs and Estimation Methods for On-Farm Research: A Simulation Study of Corn Yields at Field Scale. *Agronomy Journal*, *111*(6), 2724–2735. https://doi.org/10.2134/agronj2019.03.0142

Baars, T. (2011). Experiential Science; Towards an Integration of Implicit and Reflected Practitioner-Expert Knowledge in the Scientific Development of Organic Farming. *Journal of Agricultural and Environmental Ethics*, *24*(6), 601–628. https://doi.org/10.1007/s10806-010-9281-3

Bramley, R. G. V. & Grains Research and Development Corporation. (1999). *Designing your own on-farm experiments: how precision agriculture can help*. Kingston, A.C.T.: Grains Research & Development Corp.

Bullock, D. S., Mieno, T., & Hwang, J. (2020). The value of conducting on-farm field trials using precision agriculture technology: a theory and simulations. *Precision Agriculture*, *21*(5), 1027–1044.

Bullock, D. S., Boerngen, M., Tao, H., Maxwell, B., Luck, J. D., Shiratsuchi, L., et al. (2019). The Data-Intensive Farm Management Project: Changing Agronomic Research Through On-Farm Precision Experimentation. *Agronomy Journal*, *111*(6), 2736–2746. https://doi.org/10.2134/agronj2019.03.0165

Carton, N., Swiergiel, W., Tidåker, P., Röös, E., & Carlsson, G. (2022). On-farm experiments on cultivation of grain legumes for food – outcomes from a farmer–researcher collaboration. *Renewable Agriculture and Food Systems*, 1–11. https://doi.org/10.1017/S1742170522000102

Danko, D. M. (2012). Geospatial Metadata. In W. Kresse & D. M. Danko (Eds.), *Springer Handbook of Geographic Information* (pp. 191–244). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-72680-7_12

Hansson, S. O. (2019). Farmers' experiments and scientific methodology. *European Journal for Philosophy of Science*, *9*(3), 32. https://doi.org/10.1007/s13194-019-0255-7.

Kraaijvanger, R., & Veldkamp, A. (2015). The importance of local factors and management in determining wheat yield variability in on-farm experimentation in Tigray, northern Ethiopia. *Agriculture, ecosystems & environment*. https://doi.org/10.1016/j.agee.2015.08.003

Langford, K. (2022). Putting farmers at the centre of research to transform agriculture. *Ecos*, (285), 1–6.

Lawes, R. A., & Bramley, R. G. V. (2012). A Simple Method for the Analysis of On-Farm Strip Trials. *Agronomy Journal*, *104*(2), 371–377. https://doi.org/10.2134/agronj2011.0155

Lockeretz, W. (1995). Removing applied agricultural research from the academy. *American Journal of Alternative Agriculture*, *10*(1), 19–24. https://doi.org/10.1017/S088918930000607X

Mallarino, A. P., Wittry, D. J., Dousa, D., & Hinz, P. N. (2011). Variable-Rate Phosphorus Fertilization: On-Farm Research Methods and Evaluation for Corn and Soybean. *Precision Agriculture Proceedings,* 687–696. https://doi.org/10.2134/1999.precisionagproc4.c66

Peng, G., Lacagnina, C., Ivánová, I., Downs, R. R., Ramapriyan, H., Ganske, A., et al. (2021). *International Community Guidelines for Sharing and Reusing Quality Information of Individual Earth Science Datasets* (preprint). Open Science Framework. https://doi.org/10.31219/osf.io/xsu4p

Reetz, H. F. (2011). On-Farm Research Opportunities Through Site-Specific Management. *Precision Agriculture Proceedings,* 1173–1176. https://doi.org/10.2134/1996.precisionagproc3.c143

Steinke, J., van Etten, J., & Zelan, P. M. (2017). The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agronomy for Sustainable Development*, *37*(4), 32. https://doi.org/10.1007/s13593-017-0441-y

Sumner, M. E. (2015). Field Experimentation: Changing to Meet Current and Future Needs. In J. R. Brown (Ed.), *SSSA Special Publications* (pp. 119–131). Madison, WI, USA: Soil Science Society of America. https://doi.org/10.2136/sssaspecpub21.c11

Tanaka, T. S. T. (2021). Assessment of design and analysis frameworks for on-farm experimentation through a simulation study of wheat yield in Japan. *arXiv:2004.12741 [stat]*. http://arxiv.org/abs/2004.12741.

Thompson, L. J., Glewen, K. L., Elmore, R. W., Rees, J., Pokal, S., & Hitt, B. D. (2019). Farmers as Researchers: In-depth Interviews to Discern Participant Motivation and Impact. *Agronomy Journal*, *111*(6), 2670–2680. https://doi.org/10.2134/agronj2018.09.0626

Trevisan, R. G., Bullock, D. S., & Martin, N. F. (2021). Spatial variability of crop responses to agronomic inputs in on-farm precision experimentation. *Precision Agriculture*, *22*(2), 342–363.

Willers, J. L., Milliken, G. A., Jenkins, J. N., O'Hara, C. G., Gerard, P. D., Reynolds, D. B., et al. (2008). Defining the experimental unit for the design and analysis of site-specific experiments in commercial cotton fields. *Agricultural Systems*, *96*(1–3), 237–249. https://doi.org/10.1016/j.agsy.2007.09.003