Predictive Analytics and Data-Driven Decision Making in Small and Medium-Sized Agribusinesses in Ghana

Samuel Owusu[1]

Independent Researcher

Abstract

Predictive analytics is increasingly proposed as a means of strengthening agribusiness decisions, but the distinction between analytical capability, digital tool adoption and demonstrated decision improvement remains insufficiently examined. This study analyses Ghana's 2021 Firm Adoption of Technology survey to assess digitally supported production planning among small and medium-sized agribusiness establishments. Agriculture, livestock and food-processing records were selected using reported activity and survey size strata. Of 197 eligible establishments, 195 had a determinate planning outcome; 68 reported at least one digital planning method. A prevalence baseline, penalised logistic models and a constrained random forest were evaluated through ten repetitions of five-fold stratified cross-validation, with additional region and sector holdouts. Digital planning prevalence was 34.9% unweighted and 41.7% using survey weights. Standard software predominated, whereas specialised planning software and integrated enterprise resource planning were uncommon. No eligible respondent reported big-data analytics in marketing. Mean receiver operating characteristic area increased from 0.713 for structural predictors to 0.902 after adding digital infrastructure; the random forest reached 0.921. Removing computer counts reduced logistic discrimination to 0.822, demonstrating dependence on a proximal capability measure. These results concern classification of contemporaneous practices, not future adoption or causal improvements in profitability. The study contributes an auditable measurement and validation approach, identifying opportunities for targeted capability assessment while showing why operational deployment requires prospective outcomes, explicit decision costs and independent validation.

Keywords: *Predictive analytics, agribusiness SMEs, Ghana, digital production planning, technology adoption, cross-validation, decision support*

1. Introduction

Small and medium-sized agribusinesses make recurrent decisions about procurement, production schedules, labour allocation and customer demand. These decisions connect biological uncertainty with commercial constraints: a food processor must align perishable inputs with processing capacity, while a livestock enterprise must reconcile feed purchases with uncertain sales. Digital records could support these tasks, but access to a computer, use of

planning software and reliable predictive decision support are distinct achievements. Research on digital agriculture stresses that technological possibilities must be assessed alongside organisational conditions and uneven access to complementary resources [17, 18, 23].

Ghana provides a useful setting for examining this distinction because publicly released establishment data contain detailed business-function technology measures. The Firm Adoption of Technology (FAT) survey records technologies used for administration, production planning, procurement, marketing and other activities [9, 42]. Its function-specific design avoids assuming that every digitally connected establishment possesses equivalent analytical capability. The wider FAT research distinguishes the most advanced technology available within a business function from the technology used most intensively, demonstrating why possession alone is an incomplete description of technological upgrading [10].

Existing Ghanaian research contributes evidence on farmers' adoption of digital agricultural services and the relationship between agricultural digitalisation and livelihoods [1, 25]. However, a smallholder household and a multi-worker food-processing establishment face different information requirements and organisational arrangements. Moreover, estimating adoption associations does not establish how well a model classifies establishments outside the observations used for fitting. A predictive study must identify its target, separate training from evaluation, and explain the decisions for which the target is relevant [16, 34].

This article retains predictive analytics as its analytical method while defining the empirical outcome narrowly: reported use of at least one digital production-planning method. The outcome is a potential enabling practice for data-driven decisions, not a validated measure of decision quality. This distinction is necessary because the selected SMEs contain no affirmative responses to the marketing item explicitly mentioning big-data analytics or machine learning. Treating ordinary spreadsheet use as proof of advanced predictive analytics would therefore misrepresent the evidence.

The study asks three questions. First, how prevalent are different planning technologies among eligible Ghanaian agribusiness SMEs? Second, how much do basic digital infrastructure measures improve classification beyond structural and managerial characteristics? Third, how sensitive are apparent performance gains to geographic transfer, sector transfer and removal of computer availability? The contribution is an empirical assessment combining transparent construct definition, reproducible internal validation and a bounded interpretation of practical usefulness. It neither evaluates an implemented decision-support intervention nor estimates returns to technology investment.

2. Methodology

2.1 Study design and data provenance

A secondary, cross-sectional analysis used the public-use Ghana FAT 2021 release, reference GHA_2021_FAT_v01_M, accessed on 8 November 2026 [9]. The supplied Stata file contains 1,262 establishments and 926 variables; the CSV copy has identical dimensions and column names. The questionnaire, sector modules, codebook and methodology note were examined together. No synthetic establishments, imputed outcomes or external observations were added. Analysis was conducted at establishment level because the survey does not necessarily represent an entire multi-establishment enterprise.

The methodology describes a sampling frame derived from Ghana Statistical Service establishment censuses covering 2013–2018, telephone interviews and a 49% overall response rate [42]. Sampling stratification incorporates region, establishment size and activity. Small establishments have 5–19 workers, medium establishments 20–99, and large establishments at least 100. The methodology defines a formal private-sector universe and excludes enterprises with fewer than five employees. Accordingly, this paper does not generalise to all household farms, informal microenterprises or the full Ghanaian agricultural economy.

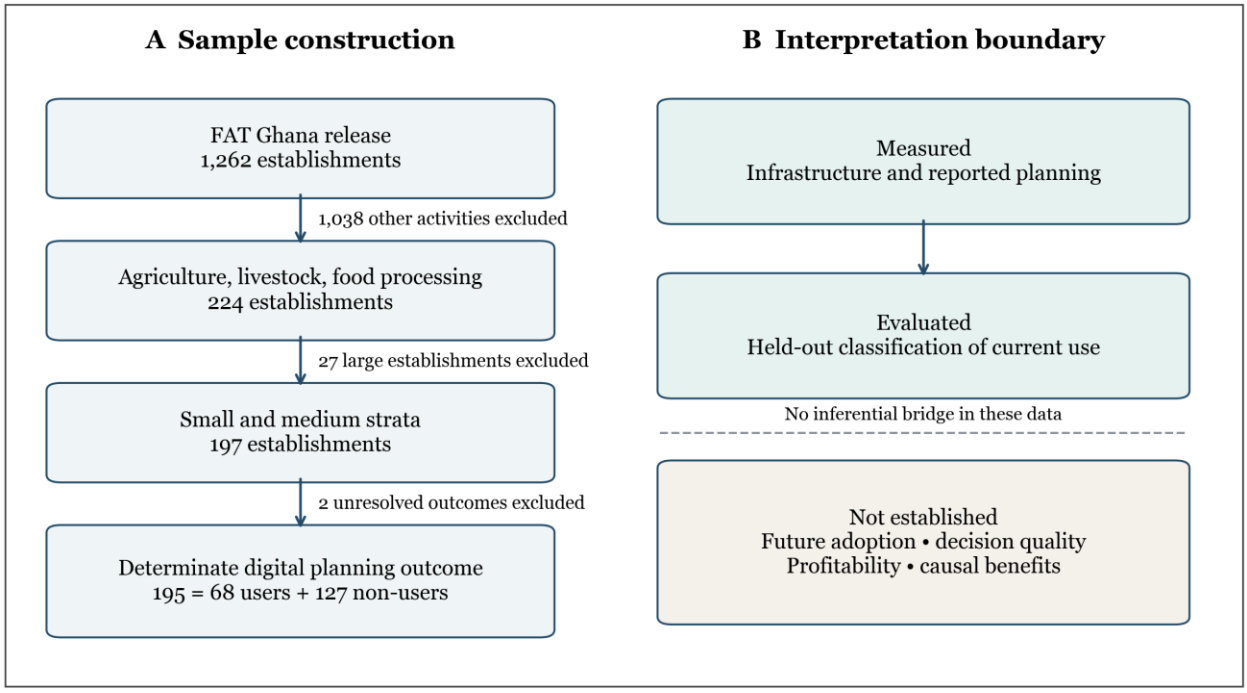
2.2 Eligibility and outcome construction

Eligibility required reported sector a5a equal to agriculture, livestock or food processing and sampling size s1b equal to Small or Medium. Reported activity was preferred to sampling activity because businesses may change activity between frame construction and interview. Of 224 records in the three reported sectors, 27 belonged to the large sampling stratum, leaving 197 eligible records: 50 agricultural, 22 livestock and 125 food-processing establishments. These are sample counts, not population totals. Figure 1 presents the selection process and the boundary between measured practices and unobserved benefits.

The binary outcome used four questions: standard computer software for planning (b9a2), mobile applications or digital platforms (b9a3), specialised planning and demand-forecast software (b9a4), and integrated enterprise resource planning (b9a5). An establishment was positive when any item equalled one. It was negative only when all four equalled two. Two establishments answered negatively to the first three items but recorded “do not know” for ERP, so their outcomes remained unresolved and they were excluded from modelling. The analytical sample therefore contained 195 records, including 68 positives and 127 negatives.

The questionnaire permits multiple technologies simultaneously. Handwritten planning was consequently analysed as a coexisting practice rather than the logical opposite of digital planning. The conditional most-used-method field b9b was not used as the outcome: it is asked only when multiple planning methods are reported. Coding its skipped responses as non-adoption would introduce a systematic measurement error. The composite measures any reported digital planning; it does not measure usage frequency, analytical sophistication or the correctness of decisions.

Figure 1. Sample construction and the boundary of inference. Selection uses reported activity and sampling size; two unresolved ERP responses prevent outcome classification. The study evaluates current planning practice, not future adoption or decision benefits.



2.3 Predictors and preprocessing

The structural model included business age, managerial sector experience, monthly power outages, manager gender, reported sector, sampling-size category, geographic stratum and manager education. Business age was calculated as 2021 minus the reported starting year. Education, sector, size and region were categorical; gender was binary. Outage frequency was transformed using log(1+x). The extended model added internet connection, desktop/laptop count and tablet/smartphone count, with device counts also transformed using log(1+x). These additions represent access to digital infrastructure rather than a composite analytics-capability score.

Outcome-defining planning variables, other planning summaries, establishment identifiers and financial outcomes were excluded from predictors. Selected predictors contained no missing values after eligibility and outcome filtering. Nevertheless, preprocessing was implemented within each training fold: numerical median imputation, scaling, categorical mode imputation and one-hot encoding with unknown-category handling. Thus held-out records did not determine preprocessing parameters. A sensitivity specification removed desktop/laptop counts while retaining internet and mobile-device measures, addressing the proximity between computer availability and computer-based planning.

Financial outcomes were deliberately not modelled. Released reference-year fields identify 2019 and 2017, while the main questionnaire contains different financial-year wording and variable mappings, including an inconsistent currency heading. Without clarification from the data producers, a harmonised profitability or sales-growth outcome would be difficult to defend. This documentation discrepancy does not directly affect the planning questions, whose labels and

response coding align. It nevertheless limits the performance claims that can be made from this release.

2.4 Models and validation

Four primary specifications were evaluated: a training-prevalence baseline, structural logistic regression, extended logistic regression and an extended random forest. Logistic models used L2 regularisation with C=1, the lbfgs solver and a maximum of 3,000 iterations. The forest used 300 trees, maximum depth four, minimum leaf size eight and square-root feature subsampling. Hyperparameters were fixed for this analysis rather than selected by maximising validation scores. No hyperparameter search or outcome-driven feature selection was performed. This design reduces tuning-related optimism but does not eliminate uncertainty from analyst specification choices [35, 38].

Ten repetitions of stratified five-fold cross-validation used seeds 20261108–20261117. Models shared identical partitions, and each establishment received one held-out prediction per repetition. Performance was computed from the pooled held-out predictions within each repetition and then averaged. Repetition standard deviations describe partition sensitivity, not confidence intervals based on independent samples. The random forest seed was fixed at 20261108. Python and scikit-learn 1.8.0 were used; the accompanying analysis scripts specify transformations, filters and model parameters.

Receiver operating characteristic area under the curve (AUROC) assessed ranking, average precision assessed positive-class retrieval, and the Brier score assessed squared probability error. Balanced accuracy used a fixed probability threshold of 0.5. Average precision was included because the classes are unequal [33]. A low Brier score alone was not treated as proof of adequate calibration: decision deployment would require larger-sample calibration assessment and an explicit assessment of errors at relevant thresholds [36]. Random forests were included as a nonlinear comparator, not presumed superior to regression [6, 8, 32].

Additional validation held out each of the six geographic strata and, separately, each of the three reported sectors. Predictions were pooled across held-out groups. These tests examine transfer beyond familiar groups within the same survey; they are not independent external validation. Entirely unseen categories were encoded as all-zero indicator blocks. Grouped validation can reveal optimism hidden by random partitions, although deliberate extrapolation changes the prediction problem itself [28]. The computer-count ablation repeated the same random and grouped evaluations after the initial models highlighted its conceptual proximity to the target.

2.5 Weighting and interpretation

Survey weights base_wt were used for descriptive adoption percentages and a secondary weighted evaluation of held-out probabilities. Models were fitted without weights, targeting discrimination within the analytical sample; weighted performance is therefore a sensitivity analysis rather than a fully design-based prediction study. The descriptive estimator was the sum of weights for affirmative responses divided by the sum of weights for valid responses. No “do not know” response was treated as “no”. Stratification, non-response adjustment and discrepancies between reported and sampled activity limit simple population interpretations [42].

The primary estimand is classification of contemporaneous reported planning practice among sampled eligible establishments. “Prediction” refers to assigning probabilities to held-out

observations, not forecasting a later year. Coefficients, model rankings and feature ablations are interpreted as associations. No significance tests of causal effects, treatment-effect estimates or claims of improved managerial decisions were derived from these observational predictions.

3. Literature Review

3.1 Digital agriculture and organisational capability

Digital agriculture encompasses information delivery, monitoring, management software and increasingly automated processes. Reviews distinguish technical development from adoption, institutional change and the distribution of benefits [17, 29, 30]. Technologies can alter labour requirements and environmental management, but their consequences depend on implementation conditions rather than an inevitable progression towards a single desirable future [11]. The supply side also matters: established agribusinesses, software providers and new entrants influence availability, interoperability and market power [5].

The organisational literature helps explain why purchasing devices may not yield reliable decisions. Gupta and George conceptualise analytics capability as a combination of resources rather than hardware alone [15]. Mikalef and colleagues link analytics capability with innovation through dynamic capabilities, while subsequent AI-capability research distinguishes technological resources from organisational capacities [26, 27]. These constructs cannot be reconstructed faithfully from a small collection of device indicators. Accordingly, this study measures infrastructure explicitly and avoids calling it a validated analytics-capability scale.

Broader digital-transformation research treats change as organisational and strategic, involving processes, capabilities and business models [39, 40, 41]. Particularly relevant is research examining big-data analytics and decision-making quality in agricultural firms in China [20]. Its direct focus on decision quality demonstrates the additional measurement needed to connect technology use to managerial outcomes. Results from that institutional setting do not establish equivalent effects in Ghana. Here, the absence of decision-quality measures requires a narrower empirical claim and motivates prospective follow-up.

3.2 Ghanaian and African evidence

Miine and colleagues studied digital agricultural service adoption among 1,199 smallholder farmers in Ghana's Bono East Region using multivariate probit and adoption-intensity models [25]. Their evidence concerns complementary services and socioeconomic or institutional correlates, rather than held-out classification of establishment planning. Addison and colleagues investigated digitalisation and smallholder livelihoods using adoption and treatment-adjustment methods [1]. Both studies establish the importance of locally grounded questions, while their household-oriented outcomes differ from the organisational practices examined here.

African reviews identify finance, skills, infrastructure and institutional arrangements as recurring constraints [3, 7, 24]. Degila and colleagues combine bibliometric analysis with a review of deployment across five West African countries, including Ghana [12]. These syntheses provide context but cannot substitute for establishment-specific measurement. Global evidence on uneven connectivity likewise cautions against treating the availability of digital services as

universal access [23]. Advisory-service research further distinguishes information transmission from effective uptake and sustained provision [2, 14].

Table 1 compares selected strands of evidence and their implications for this design. The review is purposive and problem-focused, not a systematic review with exhaustive screening. It prioritises peer-reviewed work from 2019–2026 while retaining foundational contributions on capability measurement, prediction and validation. The literature supports examining enabling conditions and adoption patterns, but does not justify assuming that a Ghanaian SME using a spreadsheet already conducts predictive modelling.

Table 1. Selected literature informing measurement and validation.

Study	Approach and contribution	Limitation and relevance
Cirera et al. (2026) [10]	Cross-country establishment analysis distinguishes advanced from intensively used technology.	Technology possession is not equivalent to routine use.
Miine et al. (2023) [25]	Ghanaian farmer survey models service adoption and intensity.	Household service adoption differs from establishment planning.
Addison et al. (2024) [1]	Ghanaian adoption and livelihood analysis uses treatment adjustment.	Does not validate this planning classifier.
Li et al. (2022) [20]	Chinese agricultural-firm survey examines analytics and decision quality.	Different setting; decision quality requires direct measurement.
Kapoor and Narayanan (2023) [16]	Methodological analysis examines leakage and reproducibility.	Supports separating training and evaluation.
Roberts et al. (2017) [28]	Methodological review examines structured cross-validation.	Supports geographic and sector holdouts.

3.3 Predictive methods and evidential standards

Agricultural machine-learning reviews cover crop, livestock, soil and water applications, often involving sensor or image data [4, 21]. Crop-yield prediction reviews and surveys of public computer-vision datasets illustrate how strongly model design depends on the outcome and observation process [22, 37]. Those prediction tasks cannot be transferred mechanically to a small cross-sectional enterprise survey. Many financial or operational decisions require time-indexed transactions, whereas FAT provides retrospective reports and contemporaneous technology descriptions.

The distinction between explanation and prediction is fundamental [34]. A model may classify digital planning accurately because computer availability is closely related to its definition, without identifying the intervention that would cause adoption. Leakage can also arise when outcome components or information unavailable at the intended decision time enter predictors [16]. This study excludes direct planning components and explicitly tests computer-count removal. Responsible innovation research adds another requirement: analytical usefulness must

be considered alongside participation, power and accountability [13, 31]. These concerns become especially relevant if models influence access to scarce support.

4. Problem Statement

The central empirical problem is a mismatch between broad claims about predictive analytics and the narrower information available in public enterprise surveys. Three gaps must be separated. The measurement gap concerns whether observed responses indicate devices, digital practices, advanced analytics or successful decisions. The validation gap concerns whether associations found in a fitted sample persist in held-out establishments and unfamiliar groups. The decision gap concerns whether better prediction improves a defined operational outcome after accounting for intervention costs and errors.

Without these distinctions, an analysis could combine ordinary spreadsheet use with specialised forecasting, report high classification accuracy driven by hardware availability, and conclude that analytics raises profitability. Each step would extend beyond the evidence. Conversely, rejecting all secondary analysis because no operational experiment is available would discard useful information about capability distribution and measurement constraints. The appropriate objective is a bounded empirical assessment of digitally supported planning, followed by a clearly identified research agenda for actual decision benefits.

The study therefore examines whether infrastructure adds predictive information beyond structural characteristics, how strong the resulting classification is under alternative validation designs, and how far the measured planning practices support claims about data-driven management. Its practical audience includes researchers and support organisations designing capability assessments. The model is not proposed as a substitute for asking managers directly about their planning practices, nor as an eligibility score for credit, grants or extension assistance.

5. Results

5.1 Sample composition and technology use

Among the 197 eligible establishments, 126 belonged to the small sampling stratum and 71 to the medium stratum. Two food-processing records had unresolved outcomes, leaving 50 agricultural, 22 livestock and 123 food-processing establishments for modelling. Positive outcomes numbered 19, seven and 42, respectively. Thirty-seven of 124 small establishments with determinate outcomes were positive, compared with 31 of 71 medium establishments. These differences are descriptive and do not establish a size effect independent of other characteristics.

Overall, 68 of 195 establishments reported some digital planning, giving an unweighted prevalence of 34.9%. The weighted estimate was 41.7%, indicating that unequal selection and non-response weights materially change the descriptive picture. The weight-based effective sample size, calculated as squared total weight divided by total squared weights, was approximately 118.1. This is a diagnostic of weight concentration, not a complete design-effect estimate. Table 2 reports adoption counts and denominators so that sparse technologies are not obscured by percentages.

Table 2. Planning practices and connectivity among eligible agribusiness SMEs.

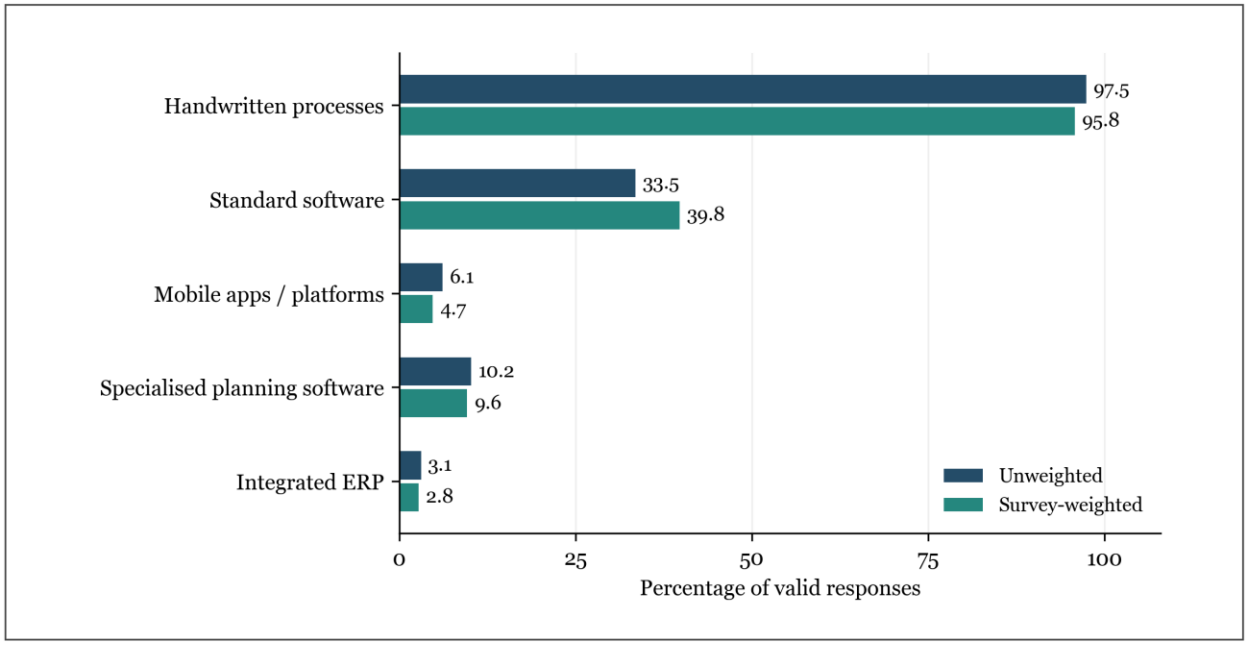
Measure	Yes / valid n	Unweighted %	Weighted %
Handwritten planning	192 / 197	97.5	95.8
Standard planning software	66 / 197	33.5	39.8
Planning apps / platforms	12 / 197	6.1	4.7
Specialised planning software	20 / 197	10.2	9.6
Integrated planning ERP	6 / 195	3.1	2.8
Internet connection	45 / 197	22.8	30.3
Marketing big-data analytics	0 / 196	0.0	0.0
Any digital planning	68 / 195	34.9	41.7

Handwritten planning was reported by 192 of 197 establishments, while standard planning software appeared in 66. Mobile applications or platforms appeared in 12, specialised planning software in 20, and integrated ERP in six of 195 valid responses. These categories overlap. Because the digital union contained 68 establishments and standard software alone contained 66, other digital methods added only two establishments beyond the standard-software users. Thus the composite outcome is empirically dominated by relatively basic computer-supported planning.

Internet connectivity was reported by 45 of 197 eligible establishments. Among the 195 modelled records, 37 of 44 connected establishments reported digital planning, compared with 31 of 151 without connectivity. This pattern does not mean internet access is necessary: spreadsheets and other software may be used offline. Computer availability was still more closely related to the target: 64 of 85 establishments with at least one desktop or laptop were positive, compared with four of 110 reporting none. Shared devices, mobile tools or reporting differences could explain those four cases; the data do not distinguish these possibilities.

The marketing analytics item had 196 valid responses, all negative, with one “do not know”. This is evidence about a specific marketing method in this sample, not proof that analytics is absent from every business function or all Ghanaian SMEs. Figure 2 compares weighted and unweighted planning adoption. Both representations show extensive handwritten practice alongside limited use of advanced planning tools, supporting an interpretation of uneven and overlapping digitalisation rather than complete replacement of manual processes.

Figure 2. Overlapping planning practices among eligible SMEs. Denominators are 197 except ERP (195). Weighted percentages use base_wt. Technologies are not mutually exclusive; their percentages must not be summed.



5.2 Predictive performance

The prevalence baseline produced a mean Brier score of 0.227. Structural logistic regression improved this to 0.202, with AUROC 0.713 and average precision 0.546. Adding internet and device measures increased mean AUROC to 0.902 and average precision to 0.807, while reducing the Brier score to 0.114. Balanced accuracy increased from 0.631 to 0.818. Table 3 reports model comparisons, including variation across repeated partitions and the two grouped validation designs.

Table 3. Held-out classification performance and transfer sensitivity.

Model	AUROC ± SD	Average precision	Brier score	Region AUROC	Sector AUROC
Prevalence baseline	0.486 ± 0.000	0.343	0.227	0.341	0.486
Structural logistic	0.713 ± 0.013	0.546	0.202	0.610	0.555
Logistic without computer counts	0.822 ± 0.013	0.733	0.155	0.763	0.770
Extended logistic	0.902 ± 0.007	0.807	0.114	0.878	0.892
Random forest	0.921 ± 0.004	0.849	0.114	0.897	0.912

The constrained forest achieved mean AUROC 0.921, average precision 0.849 and balanced accuracy 0.851. Its mean Brier score was also 0.114 after rounding, nearly identical to the extended logistic model. Thus the forest's stronger ranking did not translate into a substantial improvement in overall squared probability error. Repetition standard deviations for AUROC

were 0.013 for the structural logistic model, 0.007 for extended logistic regression and 0.004 for the forest. These values quantify split sensitivity within the available data and should not be interpreted as narrow population uncertainty intervals.

The training-prevalence baseline's pooled AUROC was 0.486 rather than exactly 0.500 because training prevalence varies slightly across stratified folds and is inversely related to each held-out fold's class proportion. Its probabilities remain constant within a fold and contain no establishment-level information. This small artefact illustrates why a pooled out-of-fold metric should be interpreted with the prediction procedure in view, rather than treated as an intrinsic property of an algorithm.

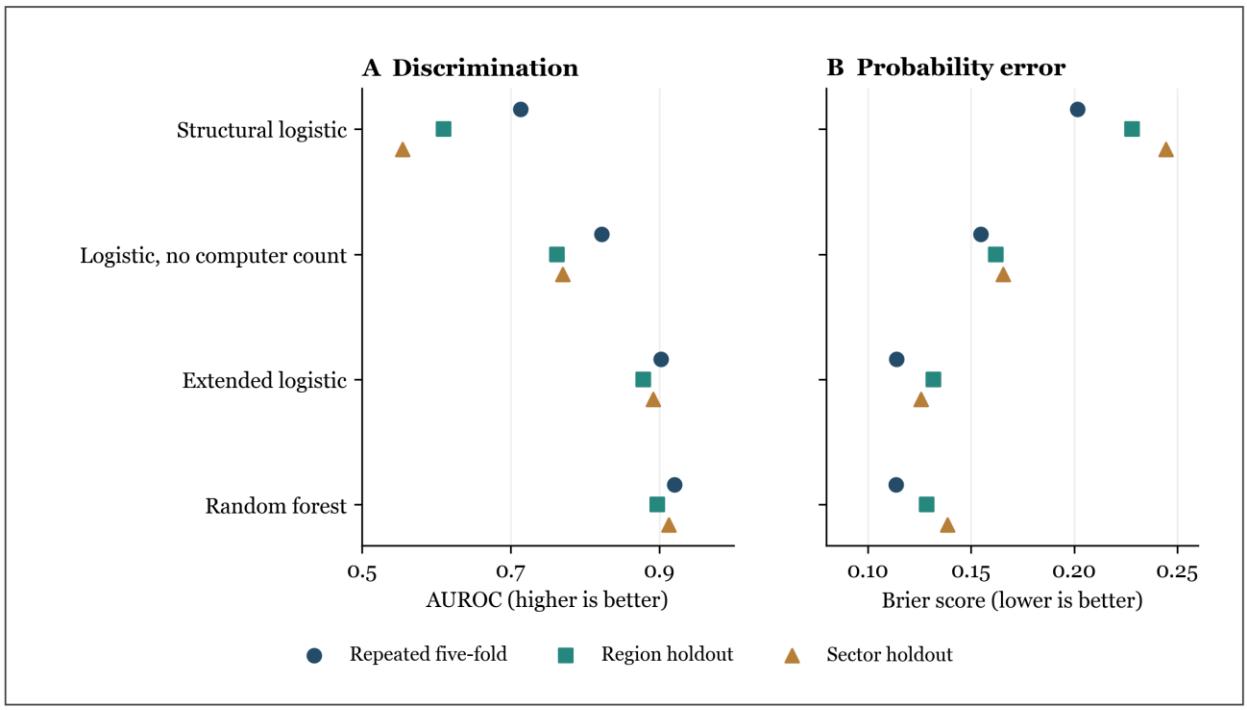
5.3 Robustness and construct proximity

Holding out entire regions reduced structural logistic AUROC to 0.610, whereas extended logistic regression and the forest retained AUROCs of 0.878 and 0.897. Their region-holdout Brier scores were 0.132 and 0.129. Holding out entire sectors produced AUROCs of 0.555, 0.892 and 0.912, respectively. However, sector-holdout Brier scores were 0.126 for extended logistic regression and 0.139 for the forest, reversing their small probability-error ranking. Model preference therefore depends partly on the evaluation setting and metric.

Removing computer counts from extended logistic regression reduced repeated-validation AUROC to 0.822 and increased Brier error to 0.155. Region- and sector-holdout AUROCs declined to 0.763 and 0.770. The ablation still outperformed the structural model, but the reduction establishes that a substantial part of classification strength depends on knowing whether establishments possess the hardware closely associated with the target. Figure 3 presents this contrast alongside the primary model results, preventing the strongest score from being detached from its measurement context.

Weighting held-out errors changed the relative picture slightly. Weighted AUROC was 0.927 for extended logistic regression and 0.945 for the forest; weighted Brier scores were 0.104 and 0.109, respectively. These estimates do not prove stronger population transportability because the models were fitted without design weights and no independent probability sample was tested. They show that the reported comparisons are sensitive to how establishments contribute to the evaluation criterion.

Figure 3. Discrimination and probability error under alternative validation designs. Repeated five-fold values are mean held-out metrics across ten repetitions. Grouped holdouts test transfer within this survey; the computer-count ablation assesses construct proximity.



6. Main Discussion

6.1 What the models establish

The clearest empirical finding is that basic digital infrastructure strongly separates current digital-planning users from non-users within this survey. This is consistent with organisational capability arguments, but substantially narrower than demonstrating that analytics improves performance [15, 26, 27]. Infrastructure may enable software use; software adoption may also motivate equipment purchase; both may reflect unmeasured finance, customer requirements or managerial preferences. Cross-sectional prediction cannot distinguish these pathways. The model's utility is therefore diagnostic and exploratory.

The ablation is particularly important. A desktop or laptop is closely connected to an outcome whose dominant component is standard computer software. Including computer counts is not direct label leakage, since the question does not itself ask whether planning software is used. Nevertheless, the conceptual relationship is sufficiently close to generate optimistic impressions of analytical sophistication. The reduction from AUROC 0.902 to 0.822 when computer counts are removed provides a more informative result than an isolated report of high predictive accuracy.

Moreover, the composite's dependence on standard software matters for interpretation. Specialised planning and demand-forecast software was reported by only 20 establishments, while the marketing analytics question recorded no affirmative SME responses. A broad label such as "predictive-analytics adoption" would conceal this distribution. The most defensible

conclusion is that some establishments possess digital planning foundations, while direct evidence of advanced analytical practice remains limited. This supports function-specific measurement rather than treating technology adoption as a single undifferentiated stage [10].

6.2 From classification to managerial decisions

A support organisation could use these findings to structure an assessment, beginning with current records, access to working devices and connectivity, then asking how information actually enters production decisions. However, directly asking the four planning questions is simpler than running a predictive model when managers are available. The classifier becomes potentially relevant only when planning responses are unavailable and predictor information has already been collected, or as a research instrument for testing the informational contribution of different capability domains.

Any operational application would require a defined action. For example, an organisation might offer introductory record-keeping support to establishments without digital planning and forecasting training to existing users. Such a programme would need to establish whether these categories predict training benefit, not merely existing technology use. A non-user might benefit greatly from assistance, whereas an established user might require little additional support. Selecting recipients according to predicted adoption would therefore risk rewarding prior advantage rather than identifying unmet need.

A prospective evaluation should record baseline practice, the offered intervention, participation, costs and subsequent decisions. Outcomes might include stock-out frequency, order fulfilment, avoidable spoilage or forecast error, provided their measurement is standardised and feasible. Repeated transaction records would permit chronological validation and comparisons against simple seasonal or managerial baselines. Randomised offers or carefully justified quasi-experimental designs could then investigate effects. These are proposed next steps, not outcomes measured or experiments completed in the present study.

6.3 Model choice and transfer

The forest's higher AUROC does not justify automatic preference over logistic regression. Their primary Brier scores were almost equal, and logistic regression had lower probability error in the sector-holdout evaluation. A relatively transparent regression model may therefore be preferable where probability communication and auditability matter [32]. Evidence from other prediction domains likewise warns against assuming that flexible algorithms necessarily improve on regression, although domain-specific results cannot be imported as agribusiness performance estimates [8].

Geographic and sector holdouts strengthen the evaluation by testing unfamiliar categories, yet remain limited. Regions share national institutions, questionnaire wording and data collection procedures; sectors share the same survey infrastructure. Strong performance in these tests does not demonstrate transportability to a later year, unregistered enterprises or another country. Digital practices can change alongside prices, connectivity and business conditions. Independent validation should therefore examine both discrimination and calibration before predictions are used for decisions [28, 36].

The weighted comparisons also illustrate competing objectives. An unweighted model comparison answers a question about the observed analytical sample; a weighted comparison emphasises establishments representing larger portions of the sampling frame. Neither

automatically corresponds to the welfare objective of an intervention. A programme targeting excluded or disadvantaged enterprises may require a different sampling design and evaluation function. Making that objective explicit is more important than selecting whichever weighting scheme produces the largest AUROC.

6.4 Implications for Ghanaian agribusiness support

The coexistence of handwritten and digital practices suggests that transition support should accommodate mixed workflows. Training could focus on reconciling paper records with spreadsheets, consistent product identifiers, documented input units and routine checks for missing transactions. These recommendations follow from the observed overlap and established organisational perspectives, rather than an estimated treatment effect [39, 40, 41]. They should be tested with managers before implementation because the survey does not describe the quality or completeness of their records.

Connectivity support should similarly reflect business functions. Offline spreadsheets may be adequate for some planning tasks, while coordination with suppliers or cloud-based services may require reliable internet. Advisory-service research highlights the difference between transmitting information and ensuring that recipients can use it effectively [2, 14]. Ghanaian studies suggest relevant institutional dimensions, including training, credit and service combinations, but the current models do not identify which package would produce the greatest benefit [1, 25].

At policy level, technology programmes should specify the problem they intend to solve and monitor distributional effects. Mission-oriented innovation research supports aligning technological investment with explicit food-system objectives [19]. Responsible-innovation scholarship and analysis of digital agriculture's politics further caution that data ownership, dependency on suppliers and participation affect who benefits [5, 13, 31]. In this setting, an appropriate use of predictive analytics would preserve managerial review and avoid converting limited survey classifications into automatic exclusion from support.

7. Challenges and Limitations

The study's most important limitation is construct validity. Digital planning is an enabling practice, not direct evidence of predictive modelling, data quality or improved decisions. The composite combines tools with different capabilities and is dominated by standard software. The specialised-software item includes demand forecasting but does not identify algorithms, forecast horizons, accuracy or managerial reliance. The marketing analytics item covers one function only. Consequently, the findings cannot estimate the prevalence or effect of all predictive-analytics applications in Ghanaian agribusiness.

Sample scope also restricts inference. The 195 modelled establishments include only 68 positive outcomes, and food processing accounts for most observations. Fewer-than-five-worker enterprises and excluded sectors are outside the methodological universe. Selection by reported activity does not reproduce the original sampling strata exactly. Survey weighting partly addresses unequal selection and non-response, but a 49% overall response rate leaves potential bias if participation depends on unobserved technology-related characteristics. Small subgroup counts prevent confident claims about comparative regional or gender performance.

The outcome is self-reported. Respondents may interpret specialised software differently, and actual use may diverge from reported possession. Two unresolved ERP responses were excluded rather than recoded; assigning both to non-use or use would bound the unweighted composite prevalence at 34.5% and 35.5% among all 197 eligible establishments. This narrow arithmetic range does not capture wider reporting error. Conditional skips in other questions and discrepancies between financial documentation and released variables further demonstrate why automatic recoding would be inappropriate.

Internal validation cannot remove all analytical optimism. The same dataset informed outcome selection, specification and evaluation, and the analysis was not preregistered. Fixed hyperparameters avoid a tuning search, but substantive design choices remain. The computer-count ablation was a diagnostic follow-up and is reported as such. Repetition standard deviations are not uncertainty intervals for new populations. Temporal validation, independent data and larger calibration samples are still needed. No causal effects, programme benefits or realised forecasting improvements can be inferred from the reported scores [16, 35, 38].

Finally, operational decision costs were unavailable. Balanced accuracy at 0.5 gives equal importance to class-specific sensitivity and specificity but does not encode the actual consequences of a missed beneficiary or unnecessary intervention. Data protection and legitimate purpose would also require assessment before any new use of establishment records. The public data support aggregate research; they do not authorise re-identification or imply that managers consented to automated decisions about finance or assistance. These limitations define the study's appropriate use rather than invalidate its descriptive contribution.

8. Conclusion

This analysis of Ghana FAT 2021 identifies a measurable foundation for data-driven planning among sampled agribusiness SMEs, while establishing clear limits on claims about predictive analytics. Of 195 establishments with determinate outcomes, 68 reported digital planning. Basic software predominated, handwritten methods remained widespread, and advanced tools were uncommon. Digital infrastructure substantially improved held-out classification, but computer-count removal materially reduced performance, showing that high discrimination partly reflects a close relationship between available hardware and the target practice.

The principal contribution is a transparent secondary-data analysis that separates observed technology use, statistical prediction and unmeasured decision benefits. Geographic and sector holdouts, probability-error metrics and a construct-proximity ablation provide a more qualified assessment than a single accuracy statistic. The results support better measurement and targeted capability assessment, not causal claims about profitability or immediate deployment of an automated support-allocation system. Future research should connect verified planning practices to time-indexed operational outcomes and evaluate whether specific interventions improve decisions for the enterprises intended to benefit.

References

[1] Addison M., Bonuedi I., Arhin A.A., Wadei B., Owusu-Addo E., Fredua Antoh E., et al. (2024). Exploring the impact of agricultural digitalization on smallholder farmers' livelihoods in Ghana. Heliyon, 10(6), e27541. https://doi.org/10.1016/j.heliyon.2024.e27541

[2] Aker J.C., Ghosh I., Burrell J. (2016). The promise (and pitfalls) of ICT for agriculture initiatives. Agricultural Economics, 47(S1), 35-48. https://doi.org/10.1111/agec.12301

[3] Arthur K.K., Bannor R.K., Masih J., Oppong-Kyeremeh H., Appiahene P. (2024). Digital innovations: Implications for African agribusinesses. Smart Agricultural Technology, 7, 100407. https://doi.org/10.1016/j.atech.2024.100407

[4] Benos L., Tagarakis A.C., Dolias G., Berruto R., Kateris D., Bochtis D. (2021). Machine Learning in Agriculture: A Comprehensive Updated Review. Sensors, 21(11), 3758. https://doi.org/10.3390/s21113758

[5] Birner R., Daum T., Pray C. (2021). Who drives the digital revolution in agriculture? A review of supply-side trends, players and challenges. Applied Economic Perspectives and Policy, 43(4), 1260-1285. https://doi.org/10.1002/aepp.13145

[6] Breiman L. (2001). Random Forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/a:1010933404324

[7] Choruma D.J., Dirwai T.L., Mutenje M.J., Mustafa M., Chimonyo V.G.P., Jacobs-Mata I., et al. (2024). Digitalisation in agriculture: A scoping review of technologies in practice, challenges, and opportunities for smallholder farmers in sub-saharan africa. Journal of Agriculture and Food Research, 18, 101286. https://doi.org/10.1016/j.jafr.2024.101286

[8] Christodoulou E., Ma J., Collins G.S., Steyerberg E.W., Verbakel J.Y., Van Calster B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of Clinical Epidemiology, 110, 12-22. https://doi.org/10.1016/j.jclinepi.2019.02.004

[9] Cirera X., Comin D., Cruz M. (2026). Ghana: The Firm Adoption of Technology (FAT) Survey, 2021. GHA_2021_FAT_v01_M [Data set]. World Bank Microdata Library. https://doi.org/10.48529/yhdd-9420 (accessed 8 November 2026).

[10] Cirera X., Comin D., Cruz M. (2026). Technology Sophistication Across Establishments. The Quarterly Journal of Economics, 141(3), 2025-2085. https://doi.org/10.1093/qje/qjag018

[11] Daum T. (2021). Farm robots: ecological utopia or dystopia?. Trends in Ecology & Evolution, 36(9), 774-777. https://doi.org/10.1016/j.tree.2021.06.002

[12] Degila J., Tognisse I.S., Honfoga A.C., Houetohossou S.C.A., Sodedji F.A.K., Avakoudjo H.G.G., et al. (2023). A Survey on Digital Agriculture in Five West African Countries. Agriculture, 13(5), 1067. https://doi.org/10.3390/agriculture13051067

[13] Eastwood C., Klerkx L., Ayre M., Dela Rue B. (2019). Managing Socio-Ethical Challenges in the Development of Smart Farming: From a Fragmented to a Comprehensive Approach for Responsible Research and Innovation. Journal of Agricultural and Environmental Ethics, 32(5-6), 741-768. https://doi.org/10.1007/s10806-017-9704-5

[14] Fabregas R., Kremer M., Schilbach F. (2019). Realizing the potential of digital development: The case of agricultural advice. Science, 366(6471), eaay3038. https://doi.org/10.1126/science.aay3038

[15] Gupta M., George J.F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049-1064. https://doi.org/10.1016/j.im.2016.07.004

[16] Kapoor S., Narayanan A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. Patterns, 4(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

[17] Klerkx L., Jakku E., Labarthe P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS: Wageningen Journal of Life Sciences, 90-91(1), 100315. https://doi.org/10.1016/j.njas.2019.100315

[18] Klerkx L., Rose D. (2020). Dealing with the game-changing technologies of Agriculture 4.0: How do we manage diversity and responsibility in food system transition pathways?. Global Food Security, 24, 100347. https://doi.org/10.1016/j.gfs.2019.100347

[19] Klerkx L., Begemann S. (2020). Supporting food systems transformation: The what, why, who, where and how of mission-oriented agricultural innovation systems. Agricultural Systems, 184, 102901. https://doi.org/10.1016/j.agsy.2020.102901

[20] Li L., Lin J., Ouyang Y., Luo X. (2022). Evaluating the impact of big data analytics usage on the decision-making quality of organizations. Technological Forecasting and Social Change, 175, 121355. https://doi.org/10.1016/j.techfore.2021.121355

[21] Liakos K., Busato P., Moshou D., Pearson S., Bochtis D. (2018). Machine Learning in Agriculture: A Review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

[22] Lu Y., Young S. (2020). A survey of public datasets for computer vision tasks in precision agriculture. Computers and Electronics in Agriculture, 178, 105760. https://doi.org/10.1016/j.compag.2020.105760

[23] Mehrabi Z., McDowell M.J., Ricciardi V., Levers C., Martinez J.D., Mehrabi N., et al. (2021). The global divide in data-driven farming. Nature Sustainability, 4(2), 154-160. https://doi.org/10.1038/s41893-020-00631-0

[24] Mhlanga D., Ndhlovu E. (2023). Digital Technology Adoption in the Agriculture Sector: Challenges and Complexities in Africa. Human Behavior and Emerging Technologies, 2023, 6951879. https://doi.org/10.1155/2023/6951879

[25] Miine L.K., Akorsu A.D., Boampong O., Bukari S. (2023). Drivers and intensity of adoption of digital agricultural services by smallholder farmers in Ghana. Heliyon, 9(12), e23023. https://doi.org/10.1016/j.heliyon.2023.e23023

[26] Mikalef P., Boura M., Lekakos G., Krogstie J. (2019). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. British Journal of Management, 30(2), 272-298. https://doi.org/10.1111/1467-8551.12343

[27] Mikalef P., Gupta M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Information & Management, 58(3), 103434. https://doi.org/10.1016/j.im.2021.103434

[28] Roberts D.R., Bahn V., Ciuti S., Boyce M.S., Elith J., Guillera-Arroita G., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913-929. https://doi.org/10.1111/ecog.02881

[29] Rose D.C., Chilvers J. (2018). Agriculture 4.0: Broadening Responsible Innovation in an Era of Smart Farming. Frontiers in Sustainable Food Systems, 2, 87. https://doi.org/10.3389/fsufs.2018.00087

[30] Rose D.C., Wheeler R., Winter M., Lobley M., Chivers C.A. (2021). Agriculture 4.0: Making it work for people, production, and the planet. Land Use Policy, 100, 104933. https://doi.org/10.1016/j.landusepol.2020.104933

[31] Rotz S., Duncan E., Small M., Botschner J., Dara R., Mosby I., et al. (2019). The Politics of Digital Agricultural Technologies: A Preliminary Review. Sociologia Ruralis, 59(2), 203-229. https://doi.org/10.1111/soru.12233

[32] Rudin C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215. https://doi.org/10.1038/s42256-019-0048-x

[33] Saito T., Rehmsmeier M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

[34] Shmueli G. (2010). To Explain or to Predict?. Statistical Science, 25(3), 289-310. https://doi.org/10.1214/10-sts330

[35] Vabalas A., Gowen E., Poliakoff E., Casson A.J. (2019). Machine learning algorithm validation with a limited sample size. PLOS ONE, 14(11), e0224365. https://doi.org/10.1371/journal.pone.0224365

[36] Van Calster B., McLernon D.J., van Smeden M., Wynants L., Steyerberg E.W. (2019). Calibration: the Achilles heel of predictive analytics. BMC Medicine, 17(1), 230. https://doi.org/10.1186/s12916-019-1466-7

[37] van Klompenburg T., Kassahun A., Catal C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709

[38] Varma S., Simon R. (2006). Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics, 7(1), 91. https://doi.org/10.1186/1471-2105-7-91

[39] Verhoef P.C., Broekhuizen T., Bart Y., Bhattacharya A., Qi Dong J., Fabian N., et al. (2021). Digital transformation: A multidisciplinary reflection and research agenda. Journal of Business Research, 122, 889-901. https://doi.org/10.1016/j.jbusres.2019.09.022

[40] Vial G. (2019). Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), 118-144. https://doi.org/10.1016/j.jsis.2019.01.003

[41] Warner K.S.R., Wäger M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. Long Range Planning, 52(3), 326-349. https://doi.org/10.1016/j.lrp.2018.12.001

[42] World Bank (n.d.). Firm-level Adoption of Technology Survey: Methodology Note—Ghana, questionnaires and codebook. Supporting documentation for GHA_2021_FAT_v01_M. https://microdata.worldbank.org/catalog/8229/related-materials (accessed 8 November 2026).