Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Candidate availability and uncertainty-aware selection shape the reliability of dairy rations under variable feed composition

Wei Liu
DeepFarm Inc
Seojin Jang
seojingud@deepfarm.net
DeepFarm Inc

Article

Keywords: dairy cattle, feed composition variability, ration formulation, candidate pool, measurement uncertainty, model-defined reliability

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11257790/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Both authors are employees of DeepFarm Inc., a company that develops precision livestock software. The authors declare no other competing interests.

Candidate availability and uncertainty-aware selection shape the reliability of dairy rations under variable feed composition

Wei Liu[1] and Seojin Jang[1*]

[1]DeepFarm Inc., Seoul, Republic of Korea.

*Corresponding author(s). E-mail(s): seojinygud@deepfarm.net;
Contributing authors: liuweinba1108@gmail.com;

Abstract

Least-cost dairy rations meeting nutrient limits at mean feed composition can fail when composition varies. We compared fixed and nutrient-informed selection rules on matched candidate pools for one reference cow using selected NASEM equations, with actual dry matter (DM) hidden during selection. With planned DM increased by 10%, adding admissible candidates reduced failure in every comparison; compression for training coverage increased it in five of six. Full-pool selection reversed these losses in all 18 test comparisons, with risks no greater than 2.155% without cottonseed and 0.825% with it. An exploratory least-cost linear program re-solved at the observed composition with the best of four DM margins failed in 8.1–9.8% of states. At higher simulated measurement noise, using posterior means alone increased failure, while adding posterior variance to the score reduced it. Energy violations became more frequent but smaller on average when they occurred. With a 20% higher microbial-protein coefficient, full-pool risks for this uncertainty-aware rule exceeded both case-specific targets, although its high-noise advantages remained. Reference-supply cases remained high-risk. These findings assume consumption of all supplied DM, use an illustrative crude-protein target and have not been field-validated.

Keywords: dairy cattle, feed composition variability, ration formulation, candidate pool, measurement uncertainty, model-defined reliability

1 Introduction

Least-cost dairy rations often place one or more nutrient supplies close to a specified limit. A ration that meets the limit at mean feed composition may miss it when composition changes, so nominal feasibility does not establish reliability. Formulation methods therefore need to be compared in terms of nominal cost and feasibility as well as the frequency and magnitude of failures under a common nutritional model [1, 2]. These criteria also apply when a formulation is revised after new information becomes available.

The relevant sources of variation depend on the decision horizon. Between-batch differences, day-to-day changes, sampling error and laboratory error each contribute to uncertainty before a ration is chosen [3–9]. A feed-table standard deviation (SD) need not describe the next batch on a particular farm, and database screening and grouping affect the reported variability [10–12]. Results obtained with a narrowed composition distribution or perfect nutrient observation require separate testing before they can support narrower uncertainty assumptions in practical feeding or be extended to noisy laboratory or near-infrared measurements.

Feed formulation has represented uncertainty explicitly for decades. Chance constraints [13] were adapted to ration formulation [14, 15] and applied in commercial feed formulation [16]. Stochastic programming and safety margins have been compared in dairy-heifer diets [17], and contemporary guidance distinguishes true composition variability from sampling and analytical variability [18]. Robust optimisation protects constraints over a declared uncertainty set [19–21]. Studies of nonlinear and multi-objective formulation also show that model representation, solver behaviour and the chosen objective affect the resulting diets [22–25]. A failed formulation grid or selection screen therefore establishes a limitation of that procedure; it does not prove that reliable feeding is impossible.

The available candidate rations and the information used to choose among them both affect selection. A constrained inventory defines a continuous formulation domain, but a practical rule may select from a finite pool of previously generated rations. Enlarging that pool can improve coverage without changing the scoring rule or the observations. Mixing admissible rations creates additional choices, although the failure probability of a mixture need not interpolate between those of its parent rations. Compressing a large pool reduces the number of options and can sacrifice coverage. Maximum-coverage and decision-focused ideas describe related effects [26, 27], but their value in ration formulation must be assessed using the nutritional event, rather than a candidate's origin or selection frequency. The timing of observation also matters: a ration fixed before testing and a rule that selects a ration after nutrient observation belong to different decision classes.

We examined whether reductions in model-defined ration failure depended on candidate availability, the information used for selection, or both. With the nutritional event held constant, we compared candidate addition, physical mixing and pool compression, then evaluated full-pool selection and the treatment of measurement uncertainty. Further analyses assessed transfer across specified dependence structures, nutrient-shortage magnitude and the high-risk reference cases. All analyses used one reference cow and the same selected nutritional model. The experiments

were developed sequentially, and each later protocol fixed its decision rules before generating its own fresh test draws. Protocol-specific departures, including historical stream reuse, are described in the Supplementary Methods. Cases A and C increased planned dry matter (DM) by 10%, without and with cottonseed, respectively; R0 and B retained the reference supply, again without and with cottonseed (Table 1). The primary variability setting, TAB, used published feed-table standard deviations. Dependence structures C0, C1 and C2 represented independent composition, nutrient dependence without DM dependence, and a transferred corn-silage DM–nutrient dependence structure, respectively. Figure 1 shows the common decision structure.

2 Results

2.1 Adding candidates reduces failure; compression removes useful options

Five candidate pools were compared: Q_0, the original admissible pool; Q_1, adding all admissible earlier price candidates; Q_2, adding training-defined physical mixtures; Q_3, Q_2 compressed by training success coverage; and Q_4, Q_3 with all Q_0 rations restored. Adding the earlier candidates reduced failure for both fixed and nutrient-informed (delta) selection in all six A/C comparisons. In case A, nutrient-informed failure fell from 3.865–4.130% to 2.310–2.445%. Physical mixing gave a further improvement in only two adjusted comparisons in A. Compression retained the selected fixed ration but raised nutrient-informed failure in five of the six comparisons (Fig. 2). These results identify the effects of the pool operations, not of the algorithms that generated the candidates.

Every rule in this comparison cost more than the nominal least-cost ration M0. Relative to M0, nutrient-informed selection on pools Q_1–Q_4 cost 1.112–1.119 times as much in A and 1.123–1.135 times in C, and the fixed rations of those pools cost 1.117–1.151 times as much in A and 1.099–1.141 times in C. The Q_0 fixed ration cost 1.169–1.180 times M0 in A and 1.228 times in C, so rules on the larger pools were cheaper than the original fixed ration but still more expensive than M0 (Table 3).

Preserving training coverage did not preserve test performance. In every comparison, compression added 63–91 states per 20,000 in which all candidate rations failed. It therefore raised the failure floor under complete information. Failures above this floor fell in two comparisons and rose in four. Because the floor assumes access to hidden actual DM, failures above it combine information limits with any additional loss from the implemented rule. The compressed rules nevertheless met the adjusted 5% (A) and 2% (C) criteria. Restoring the original rations (Q_4) did not yield a resolved further gain. Under narrower variability, the nutrient-informed rule on Q_0 had no observed failures in the six A/C tests, compared with 3–22 per 20,000 states for compressed pools (Supplementary Methods S31).

2.2 The full pool reverses the loss from compression

Keeping all 599 candidate rations in A and 978 in C lowered failure relative to the compressed pools in all 18 fresh paired comparisons, to 1.765–2.155% in A and 0.520–0.825% in C (maximum adjusted upper confidence bounds 2.6437% and 1.1441%). Every test met its 5% (A) or 2% (C) criterion; the stricter 1% criterion held in all three C replicates under C2 but in only two under each of C0 and C1.

These ranges and the Q_2 risks in the attribution experiment come from different test populations. The attribution experiment evaluated the same full pool on its own C0 states; the coverage protocol used fresh C0, C1 and C2 states. In the attribution experiment, risks were 1.950–2.130% in A and 0.665–0.795% in C. Only two of the three C replicates met the adjusted 1% criterion; the remaining replicate had an upper confidence bound of 1.1090%.

Partial scoring made the same choices as full scoring without pruning, while evaluating energy for 47.73–50.16% of candidate–state combinations. This percentage measures energy evaluations rather than runtime (Fig. 3; Supplementary Methods S35).

In an exploratory control using three independent C0 development streams per case, least-cost reformulation at the observed composition had higher failure rates than the nutrient-informed rules that selected from a pool. Re-solving the least-cost linear program (LP) with observed nutrient concentrations and no margin gave failure rates of 59.0–60.0% in A and C, compared with 87.2–87.9% for the nominal ration M0 on the same states. Mean cost was only 0.6–0.7% above M0. Binding constraints had zero planning margin, while actual DM fractions remained unobserved. The crude-protein (CP) and linear-energy constraints were violated in about half of the states, as under M0. The reference-chain energy constraint, which contributes to the failure event, was violated in 26–27% of states.

Of the four DM margins tested (2.5, 5, 10 and 15%), 5% gave the lowest failure rates: 9.6–9.8% in A and 8.1–8.6% in C (upper confidence bounds 10.5–10.8% and 9.0–9.5%). The 5% re-solve cost 1.119–1.121 times M0, close to the ratios of 1.112–1.135 for nutrient-informed selection on pools Q_1–Q_4 (Table 3), although costs were not matched between policies. The full-pool rules failed in 1.765–2.155% (A) and 0.520–0.825% (C) of their own test states drawn from C0, C1 and C2. Larger margins made the LP infeasible in most states, which then used M0 as a fallback. The LP also used a different cost objective and a linear energy constraint, so the control supports conclusions only about these specific policies (Supplementary S37).

2.3 Propagating posterior variance improves selection at the higher noise level

On the compressed Q_3 pool at the higher assumed noise level, using posterior means alone (the posterior-mean rule) raised failure relative to the raw-observation rule in all six paired comparisons. Adding posterior nutrient variance to the same means (the uncertainty-aware rule) lowered failure in all six: by 4.025–6.445 percentage points relative to the posterior-mean rule and 1.625–2.530 points relative to the raw-observation rule. At the lower noise level, the corresponding adjusted intervals included zero.

Under the specified working prior and first-order score, the high-noise improvement was associated with propagating uncertainty rather than correcting means alone.

With the full pool, uncertainty-aware selection met the adjusted 5% criterion in all three replicates of both A and C. Observed risks in C were below 2%, but adjusted upper confidence bounds of 2.039–2.123% did not meet that stricter criterion. The full pool also improved on compression in five of six comparisons at each observation level (ideal, lower and higher noise); the remaining intervals included zero. Measurement error was hypothetical, independent and simulated only under C0. Raw-observation and posterior-mean scores were tested only on Q_3. The design therefore does not establish whether uncertainty-aware selection outperforms raw selection on the same full Q_2 pool, or whether pool coverage and measurement treatment interact (Fig. 3).

2.4 Frozen rules transfer across dependence structures; R0 and B stay high-risk

The forward experiment froze finite-node rules under C0 and tested them under C0, C1 and C2 with five information sets: no observation (I0), true DM (I1), noisy DM (I2, two noise levels), ideal nutrients (I3) and noisy nutrients (I4, two noise levels). We discuss I0, I1 and I3 here. Relative to no observation, ideal nutrients lowered failure in all 18 adjusted pairs. True DM lowered failure in 15 pairs, with three unresolved. The finite-node rule itself performed worse than ideal delta on the same pool in 14 of 18 cells, so these comparisons address the effect of information rather than the choice of score. Without retraining, the nutrient-informed rule frozen under C0 met A's adjusted 5% target in every C0, C1 and C2 replicate, and C's 2% target in all C1 and C2 replicates but only two C0 replicates.

In the reverse experiment, rules trained on the expanded pool under the historical C2 structure were tested on new C2 and C0 states. Nutrient-informed selection improved on the frozen fixed rule from the same pool in all 12 pairs and met A's 5% and C's 2% targets. No C reverse record met the adjusted 1% criterion. The uncertainty-aware rule was tested only under C0, so its ability to transfer remains unestablished (Supplementary Methods S33).

In the reference-supply cases, fixed rules using rations generated during training in the continuous domain lowered failure relative to nutrient-informed selection on the parent pools: from 90.255–90.635% to 71.750–72.860% in R0 and from 93.175–93.415% to 64.125–64.380% in B. None of the 24 risks met the 20% target. Lower confidence bounds on the fraction of states in which all candidate rations failed also excluded that target within each new finite pool (Supplementary S36).

2.5 Failure frequency, shortage magnitude and sensitivity to microbial protein

The full pool reduced mean energy shortage. At high noise, uncertainty-aware selection made energy shortages more frequent but smaller on average when they occurred. Relative to Q_3, full Q_2 reduced the mean shortage of net energy for lactation (NEL) over all states with defined magnitudes in all 18 coverage comparisons and all six measurement comparisons at each observation level. Changes in other constraints varied in

direction. On Q_3, the uncertainty-aware rule increased this mean relative to both raw observations and posterior means alone in all six comparisons at each noise level. At high noise, it increased NEL violation incidence relative to raw selection but lowered the mean and 95th-percentile shortages conditional on violation in all six comparisons. In A replicate 0, incidence rose from 2.700% to 3.530% of all test states, while the mean shortage conditional on violation fell from 1.138% to 1.099% of the energy target. The increase in incidence outweighed the smaller conditional mean. Because the violating states differed among policies, these conditional summaries do not imply smaller shortages in every matched state. CP, Ca, P and three of the five neutral detergent fibre (NDF)/starch rules (T1, T3, T5) had lower mean shortages relative to both controls in every high-noise comparison (Fig. 4; Supplementary Methods S35).

Absolute target attainment was more sensitive to the selected energy chain's microbial-protein coefficient than were the relative improvements. The chain differs structurally from the complete calculation of the National Academies of Sciences, Engineering, and Medicine (NASEM; component checks in Supplementary Methods S29). With the original rations held fixed, a 20% increase in this coefficient gave ideal full-pool risks of 2.700–3.175% in A and 0.895–1.435% in C. High-noise full-pool uncertainty-aware risks rose to 5.060–5.105% and 2.345–2.580%, exceeding the 5% and 2% thresholds even as point estimates. Across all five coefficient levels, full pools had lower empirical failure than compressed pools in every coverage and measurement comparison. High-noise Q_3 uncertainty-aware selection also retained lower failure than both controls in all six comparisons, although some low-noise differences reversed. These post-hoc analyses use hypothetical coefficient levels and provide no new paired significance claims or prospective risk certificates (Supplementary Methods S35).

3 Discussion

Candidate availability and the treatment of uncertainty after observation both affected how much selection reduced failure under variable feed composition. Candidate availability produced the more consistent improvement: adding admissible rations reduced failure, physical mixing added little, and compression sometimes removed useful choices. Scoring corrected concentrations alone could mis-rank rations when substantial uncertainty remained; including posterior variance improved selection at the higher assumed noise level. Each result applies to the comparison with its own matched control.

The failure rate under complete information was computed using actual DM, which the selector never observed. The gap to this lower bound therefore combines the effect of hidden DM with any excess loss from the implemented rule; it cannot be attributed solely to scoring error.

Posterior means are point estimates and do not describe the remaining posterior dispersion. Adding nutrient variance to the same means changed the ranking of candidate rations and improved the higher-noise comparisons. The score remains an approximation because the working prior differed from the truth distribution and covariance and nonlinear corrections were omitted. The adverse result for posterior means alone therefore does not establish a general limitation of mean correction. Two

properties further complicate the rankings. Summed marginal proxies approximate the expected number of failed constraints rather than the probability of any failure, so overlaps between failures can reverse rankings. Added variance also raises a normal-tail proxy at a positive tolerance-adjusted margin but lowers it at a negative margin.

Failure frequency and shortage magnitude describe different aspects of a decision. For $Z \sim N(0,1)$, a binding margin $M = \sigma Z$ has $\Pr(M < 0) = 1/2$ at every positive σ, while $\mathrm{E}[(-M)_+] = \sigma \mathrm{E}[(-Z)_+]$ decreases as variability decreases. Conversely, a low failure probability can conceal rare but substantial shortfalls. Compared with raw observation on the compressed pool at high noise, the uncertainty-aware rule produced more frequent energy violations, but the mean and 95th percentile of shortage conditional on violation decreased. Mean shortage over all defined states combines frequency with conditional size; it increased because of the higher incidence. The conditional summaries describe different subsets of violating states, and changes in maximum shortage had mixed directions.

The full generated pool had lower failure than the compressed pools in all 18 coverage comparisons. Retaining the full pool was therefore preferable to selecting from a short list in these simulations. Even the larger finite pools had limits. Candidates generated during training reduced failure in the reference-supply cases, but the frozen pools still left many states in which every candidate ration failed. Additional information cannot improve selection in those states. These bounds apply only to the tested pools, just as the empty finite grids of the original methods apply only to those grids; neither establishes that every ration in the continuous domain fails.

In the exploratory LP control, least-cost re-formulation at the observed composition did not reach the low failure rates of the nutrient-informed rules that selected from candidate pools. Without a margin, the re-solved ration failed in about three states out of five, consistent with binding rows at zero nominal margin while DM fractions remained unobserved. The best of the four DM margins tested lowered failure to 8.1–9.8% on the control's own development states. Mean cost was similar to that of the pool-based rules, but costs were not matched. The approaches also differed in their objectives and nutritional calculations. The LP used a cost objective, a uniform DM margin and a linear energy row. The pool-based rules selected among candidates from several generators using a score over the nine rows of the failure event. This control cannot identify which difference accounts for the failure-rate gap.

The 10% increase in planned DM in cases A and C tests sensitivity to supply by representing a cow that eats more than the population prediction. The NASEM Eq. 2-1 prediction comes from a model with a cross-validated root mean squared error of prediction of 2.89 kg/d on a database with a mean intake of 24 kg/d [28], so an intake 10% above the prediction lies within one prediction error. This population error provides no intake interval for the reference cow and does not establish that higher intake would be sustained on every ration. Because the simulation assumes consumption of all supplied DM, cases A and C compare planned-supply assumptions, not animal responses to extra feed. At the reference supply, case B increased energy density by including cottonseed but did not restore reliability. The constraint-pricing analysis of the original problem (Supplementary S16–S20) found almost no headroom on the linear energy row around the nominal least-cost solution at the reference supply.

The continuous rations generated during training in Supplementary S36 lowered the R0 failure rate by only about one fifth, leaving it far above every target. Both results, like the finite-pool bounds, concern only the procedures tested.

The arithmetic of the selected nutritional calculations was checked, but consistency between components does not establish biological adequacy. Structural differences from complete energy calculations and the illustrative CP target remain part of the model definition. Changing the microbial-protein coefficient fMCP while holding rations fixed preserved the high-noise advantages, and relative gains were more stable than attainment of absolute targets. The response of a retrained policy remains unknown. The study used one reference cow, a restricted inventory and selected nutritional equations; batches on a particular farm may vary differently from the feed-table distributions. Measurement errors were hypothetical and independent, and the uncertainty-aware rule was tested only under C0.

Within the selected model, full-pool selection avoided the losses caused by compression, and variance propagation improved selection under substantial assumed measurement noise. Lower failure was accompanied by higher feed cost relative to M0: nutrient-informed selection on pools Q_1–Q_4 cost 1.112–1.135 times M0 under the stated prices. This expenditure difference does not estimate a price of reliability. Some targets remained unmet: the high-noise C tests missed the adjusted 2% criterion, the coefficient sensitivity raised high-noise risks above both case-specific targets, and the reference-supply cases remained high-risk. Interpreting a low simulated failure rate as a reliable feeding decision therefore requires consideration of the available candidate rations, the measurement model, and both the frequency and magnitude of the remaining shortages.

4 Methods

4.1 Cases and selected nutritional model

The reference animal was the mature Holstein cow at 200 days in milk specified in NASEM Table 21-1. Milk composition, pregnancy status and body reserves followed Tables 21-1 and 21-3 [29]; body condition score was 3.0, with thermoneutral confinement and a total mixed ration. Inventory R0 comprised corn silage, mid-maturity legume hay, dry ground corn, solvent-extracted soybean and canola meals, soybean hulls, limestone and dicalcium phosphate. The two mineral ingredients had fixed compositions. The inventory had one principal starch source and no added fat.

Case A increased planned DM supply to 110% of the animal-factor intake predicted by NASEM Eq. 2-1. Case B added whole linted cottonseed, capped at 10% of planned DM as a study assumption informed by NASEM's discussion of gossypol. Case C combined both changes (Table 1). Nutritional targets were the same in all cases, and the linear energy constraint used in planning was recalculated for each supply level. All supplied DM was assumed to be consumed; planned supply therefore does not predict voluntary intake. The price scenario (US Midwest, September 2026) used recorded source values and declared assumptions [30–34]. It excluded freight, processing, shrink and premix. FEEDVAL mineral entries were assumptions used for comparison rather than price quotes (Supplementary S1).

The main nutritional event covered nine constraints: NEL, CP, absorbed calcium and phosphorus supply, and five NDF/starch rules from NASEM Table 5-1. The CP lower bound was the illustrative intake in Table 21-3. It served as a case target rather than a test of adequacy for metabolisable protein, essential amino acids or rumen-degradable protein. Other minerals, vitamins and physically effective fibre were not modelled. The study's upper limits on CP and ether extract applied only during planning and were outside the probability event. Equality of planned DM supply, inclusion caps and the starch-source premise were structural constraints.

Evaluation recalculated the selected Chapter 3 energy chain at realised supplied DM, retaining the original fixed approximation for microbial protein. This was not an evaluation of the full NASEM model. Because Table 5-1 assumes that starch comes predominantly from dry ground corn, corn had to supply at least half of planned starch, a threshold set for this study. Verdicts were classified as unknown when this premise failed. The model represents this premise alone; it does not represent the processing and particle-size effects studied elsewhere [35, 36]. The original constraints are documented in Supplementary S1–S5 and S16.

4.2 Composition distributions and decision information

DM, CP, NDF, starch, ether extract (EE), ash, calcium and phosphorus accounted for 48 random composition cells in R0/A and 56 in B/C. Each cell used a truncated-normal parent fitted to its target mean and SD, or a moment-matched scaled beta when no such parent existed. TAB used the SDs in NASEM Table 19-1. SRC12 and SRC14 narrowed selected corn-silage cells using temporal ratios for the same component. MO12 and NARROW also applied a corn-silage DM ratio to other cells [5]. All five distributions were specified for sensitivity analysis rather than estimated from farm populations; NARROW generally assumes lower variability.

The second extension combined TAB and NARROW with three dependence structures: independence (C0); historical nutrient correlations with independent DM (C1); and the complete historical corn-silage DM/CP/NDF/starch block together with the other nutrient pairs in C1 (C2). Policies were trained separately for each structure. Marginal distributions, targets, planned supply, the Table 5-1 premise and the fixed energy approximation were unchanged.

A candidate ration was a daily as-fed vector q_i, fixed or selected from a frozen candidate pool. For nutrient concentration a_{ij}, actual supplied-DM fraction d_i and as-fed unit price p_i,

$$D(q,d)=\sum_i q_i d_i, \qquad N_j(q,\theta,d)=\sum_i q_i d_i a_{ij}, \qquad C(q)=\sum_i p_i q_i.$$

Planning used the DM estimates $\hat{d}_i$ available at decision time. Concentration constraints used N_j/D, and energy was expressed in Mcal/day (4.184 MJ/Mcal). The ideal nutrient-informed (delta) rule observes the modelled non-DM concentrations before choosing q, while actual test DM remains hidden. A fixed rule ignores these observations and uses one ration selected during training for each cell (Fig. 1).

4.3 Candidate pools and selection rules

The original study compared the nominal least-cost LP ration, M0, with coefficient-shift safety margins, joint chance-constrained sample-average approximation, and box and budgeted robust programs [21, 37, 38]. These methods were evaluated on finite grids and solved with HiGHS through SciPy [39, 40]. Infeasible solutions, timeouts and empty outcomes were retained (Supplementary S5–S20). Results from these finite grids do not establish infeasibility in the continuous domain.

We evaluated a finite candidate pool because formulation software can generate and vet rations offline, then select among them automatically at feeding time. This approach keeps candidates within the structural domain and allows the selection rule to be audited. Pool size is a design choice; the compression comparisons quantify its effect on performance. The LP re-solve control described below evaluates an alternative: least-cost reformulation at the observed composition. The second extension compared seven predefined rules using 4,000 training and 20,000 test states. The original admissible pools contained 8/332/1/353 rations in R0/A/B/C. Adding every admissible ration from 15 earlier raw price-candidate files per case, without filtering by earlier scores, increased these pools to 8/515/2/894.

These files contained the fifteen candidate sets frozen for the same case in the earlier price extension: one set for each of the five composition-distribution settings (TAB, SRC12, SRC14, MO12 and NARROW) in each of three replicates. They were read without using that extension's validation choices, risk estimates or test results. Each set combined the original candidates for that setting (the nominal LP, safety-margin, joint chance-constrained sample-average, and box and budgeted robust programs on their original grids), rations re-solved from the same grids on 128 training states at α of 0.01, 0.05 and 0.10, and the terminal rations from the original local sequential-LP search. For B and C, the re-solves and the search were repeated at the low, midpoint and high registered cottonseed prices.

A ration was admissible if it was finite and non-negative, satisfied every structural constraint of the case (planned-DM equality, inclusion caps and the dry-ground-corn starch-share constraint) at nominal composition and decision-time DM estimates, and met the domain premise of Table 5-1 at the planning stage. This check did not include the probabilistic nutrient constraints. Of the 183 rations added in A and 541 in C, 104 and 302 came from the local search, and 79 and 239 from the re-solved grids. The Q_1 comparison therefore evaluates enlargement with candidates from several sources rather than the contribution of a single generator.

Greedy success coverage selects the ration that passes the most training states not already covered [26]. It selected up to eight rations for pairwise mixing at weights 0.25, 0.50 and 0.75, with each mixture evaluated through the reference chain. The same procedure subsequently compressed the enlarged pool to at most 64 rations, stopping when no further coverage was gained. Pools and rules were frozen before test sampling. The Supplementary Information and data files use "menu", "recipe" and "legal" to mean pool, ration and admissible, respectively.

The attribution protocol used dependence structure C0 under TAB and NARROW, with three independently trained replicates for A/C and TAB negative controls for R0/B. It compared five pools: Q_0, the original admissible pool; Q_1, with every

admissible earlier price candidate added; Q_2, with physical mixtures added; Q_3, formed by compressing Q_2; and $Q_4 = Q_0 \cup Q_3$, restoring all original rations without a size cap. Thus, Q_4 tests restoration rather than serving as a control with the same 64-ration capacity. Within each cell, all pools shared the targets, structural domain, planned supply, observations, scoring rule and the 4,000 training and 20,000 test states. These comparisons identify the effects of the pool operations, not of the algorithms that generated the earlier candidates.

Each pool had a fixed rule that prioritised the lowest training failure count, with ties broken first by cost and then by hash. Its nutrient-informed rule ranked margins for the observed state using normal-tail approximations and training DM moments. If all scores were nonfinite, it used the fixed ration as a fallback.

A separate benchmark used the original local sequential-LP search in the continuous domain defined by the structural and box constraints, starting from Q_0 or Q_2. It used the first 128 training states, three 20-second rounds and a cost objective, whereas the fixed selector used 4,000 states.

4.4 Risk inference and finite-pool limits

The second extension also tested two integration-based scores on the compressed pool: a joint-event score using 64 scrambled Sobol nodes (Joint64) and a partly analytic Rao–Blackwellised score using 16 nodes (RB16) [41, 42]. Neither showed a resolved benefit over delta on the same pool in the A/C comparisons. Conditioning does not bound the error from a fixed set of nodes [43]. Derivations and adverse outcomes are documented in archive S23–S26.

Failure was defined as $L = V \vee U$, where V indicates a known, assessable reference violation and U an unresolved verdict in the absence of such a violation. Unknown verdicts therefore counted as failures. States with unknown verdicts or failed domain premises remained in the risk denominator. The descriptive bracket $[n_V/n, (n_V + n_U)/n]$ represents classification uncertainty. Targets such as 5% or 2% were research criteria rather than animal-safety standards.

Conditional on each frozen policy and its simulation distribution, test outcomes were treated as independent Bernoulli observations. Rates used one-sided Clopper–Pearson bounds [44] with 0.05/10,000 per rate claim in later protocols. For paired policies on the same states, n_{10} and n_{01} count failures unique to the policy and to its comparator, and the difference $(n_{10} - n_{01})/n$ has the conservative interval

$$[L_\gamma(n_{10}) - U_\gamma(n_{01}),\ U_\gamma(n_{10}) - L_\gamma(n_{01})], \qquad \gamma = \frac{0.05}{4 \times 10{,}000}.$$

Here $L_\gamma(k)$ and $U_\gamma(k)$ are one-sided Clopper–Pearson bounds for k events in the same n states, each with tail error γ. The rate, paired, pool-bound and exploratory families each received a separate allocation of $\alpha = 0.05$, with a bound of 10,000 claims. These allocations do not support an overall joint 95% claim. Independently trained replicates were not pooled into a 60,000-state certificate.

For a frozen pool Q, $h_Q(S) = \min_{q \in Q} L(q, S)$ marks states where every ration fails, even with full actual-DM information. Let Y be the observations available to an

implementable rule $\pi(Y) \in Q$. The statewise identity is

$$L(\pi(Y), S) = h_Q(S) + [1 - h_Q(S)]L(\pi(Y), S).$$

Its second term, failure despite a successful full-state option, may reflect selection error or an option that depends on hidden DM. Define the full-state floor and optimal observation-limited risk by

$$R_Q^{\text{full}} = \mathrm{E}[h_Q(S)], \qquad R_Q^Y = \mathrm{E}\left[\min_{q \in Q} \mathrm{E}\{L(q, S) \mid Y\}\right].$$

Then for $R(\pi) = \mathrm{E}[L(\pi(Y), S)]$,

$$R(\pi) = R_Q^{\text{full}} + \underbrace{(R_Q^Y - R_Q^{\text{full}})}_{\text{information restriction}} + \underbrace{(R(\pi) - R_Q^Y)}_{\text{excess over the optimal observable rule}}.$$

For a fixed trained replicate, all three terms are nonnegative. The counts identify the first term and the sum of the last two. Separating the last two would require the optimal rule based on the available observations, which was not computed here. Simultaneous lower bounds on R_Q^{full} apply only to the specified pool, not to larger pools or the continuous domain.

4.5 Full-pool partial-score evaluation

Full-pool selection retained every Q_2 ration. The original delta score sums nine nonnegative marginal tail proxies, $S(q) = \sum_{j=1}^{9} p_j(q)$. For exact probabilities,

$$\sum_j P(F_j \mid Y) = \mathrm{E}\left[\sum_j \mathbf{1}_{F_j} \mid Y\right] \geq P\left(\bigcup_j F_j \mid Y\right).$$

For exact probabilities, this sum is the expected number of failed constraints and can exceed the probability of any failure. Differences in failure overlap among candidates can therefore change their rankings. The eight non-energy terms give $B(q) \leq S(q)$. Once an evaluated ration has score u, any ration with $B(q) > u$ cannot minimise the proxy, so its energy calculation can be skipped. The compressed pool Q_3 was used only for the initial evaluations. Six C0-trained A/C pools were tested under C0/C1/C2 on independent streams of 20,000 states, comparing full Q_2 with compressed Q_3 within each cell (Supplementary Methods S35).

4.6 Measurement uncertainty and uncertainty-aware scores

Observed concentrations were treated as measurements of latent model inputs subject to error [45]. Each of the six previously trained C0/TAB A/C parent replicates received a fresh test of 20,000 states and an independent noise stream, without retraining or selection based on the test results. Observations of the seven mapped nutrient

fractions were $Y = \min(1, \max(0, X + \lambda s Z))$, $\lambda \in \{0.1, 0.5\}$, with the training SD s held fixed. All noise rules within a cell used the same saved standard normal draws; latent values, missing inputs and training point masses were unchanged.

Each field had an independent Gaussian working prior, truncated to $[0, 1]$, with location and scale fixed from training. This prior differed from the distribution used to generate latent values. For interior observations, conjugate Gaussian moments were calculated before truncation. With $\sigma = \lambda s$, the endpoint likelihoods $\Phi(-x/\sigma)$ and $\Phi((x-1)/\sigma)$ explicitly accounted for censoring. Margins were evaluated at posterior means. The uncertainty-aware rule used the first-order variance

$$V_r = \sum_i \left(\frac{\partial m_r}{\partial d_i}\right)^2 s_{d_i}^2 + \sum_{i,j} \left(\frac{\partial m_r}{\partial \theta_{ij}}\right)^2 \mathrm{Var}(\theta_{ij} \mid Y),$$

while retaining the original constraint directions, units, denominators and tolerances. Covariance and higher-order terms were omitted. Energy gradients used sampled CP, NDF, starch and ash, with fatty acids and lignin held fixed; sampled EE contributed no smooth variance (Supplementary Methods S29). Ten predefined rules compared raw-observation, posterior-mean and uncertainty-aware scores and coverage by Q_2/Q_3. Every Q_2 candidate was scored. For raw signed margin m and constraint tolerance τ, the effective margin is $a = m + \tau$. At variance $V > 0$, the implemented proxy is $p(V) = \Phi(-a/\sqrt{V})$, with

$$\frac{\partial p}{\partial V} = \frac{a}{2V^{3/2}} \phi(a/\sqrt{V}).$$

At fixed a, the derivative is positive for $a > 0$, negative for $a < 0$ and zero for $a = 0$. Adding variance can therefore reorder candidates, but this sign pattern does not explain the observed NEL trade-off (Supplementary S35).

4.7 Observed-composition LP re-solve control

As an exploratory control, we also reformulated the ration at the observed composition using a least-cost LP. For every development state of cases A and C under TAB with independent cells, the nominal least-cost LP was re-solved using the observed, error-free nutrient concentrations in all nine nutrient constraints. This provided the same information as the ideal nutrient-informed rule. The energy constraint was re-linearised at the planned DM supply. DM fractions remained at their planned values, all structural constraints were unchanged and as-fed cost was minimised. The resulting ration was evaluated on the realised state using the same reference event, with the nominal ration M0 evaluated on the same states as a paired reference.

A second variant added a safety margin for hidden DM using the original safety-margin method's explicit-spread interface. With composition spread set to zero, each DM fraction in a nutrient constraint was shifted by a factor $1 \pm \kappa$ in the direction that tightened that constraint, with κ in $\{0.025, 0.05, 0.10, 0.15\}$. If the LP with this margin was infeasible, M0 was used and the state was counted as a fallback. The control used three independent development streams of 20,000 states per case, outside the frozen protocols and their reserved roots. It used the same one-sided Clopper–Pearson bound (per-claim error 0.05/10,000) and paired counts, but remained an exploratory,

unregistered comparison outside the predefined comparison sets. Its code and public results are available in the repository and in Supplementary S37.

4.8 Information and directional policy transfer

The information experiment applied the C0-trained compressed pool and nuisance model to C0/C1/C2 targets without modification. A common loss for the complete event used 64 frozen empirical nutrient/DM node pairs from C0 under five information rules: no observation (I0); true DM with nutrients integrated (I1); noisy DM (I2); ideal nutrients with DM integrated (I3); and noisy nutrients (I4). Independent Gaussian noise had an SD of 0.1 or 0.5 times each field's training SD. Only observations were censored to $[0, 1]$, then treated as exact inputs. Delta on the same pool and the 4,000-state fixed rule served as controls. Each frozen replicate was tested on 20,000 fresh states per target.

A reverse experiment transferred the fixed and nutrient-informed rules trained on the historical C2 expanded pool to new C2 and C0 tests without modification. Comparisons were paired within each target's states. Because the two experiments used different pools, forward and reverse results could not be paired (Supplementary S33).

4.9 Continuous candidate generation in high-risk cases

Candidate generation for R0 and B used only the 4,000-state training samples. The parent pools were the C0 Q_3 attribution pools rather than the original Q_0 pools (Supplementary Methods S36). Two non-energy max-slack LP starting points (training means or all rows) and the parent fixed ration seeded local SLSQP searches. These searches minimised mean smoothed maximum or squared full-reference deficits over 512 fixed, equally spaced training states. They retained the original structural and box constraints, requirements and planned DM, using smoothing temperatures of 0.02 and 0.05, at most 100 iterations and a tolerance of 10^{-9}. No global optimum is claimed. Every returned ration was evaluated with the canonical checker on all 4,000 states, and all parent rations were retained in the pool. Six C0/TAB tests of 20,000 states compared parent-pool delta, admissible-parent delta, new fixed and new delta (exploratory comparisons and pool floors: Supplementary S36).

4.10 Component comparisons, shortage magnitude and implementation checks

A component audit compared the selected chain with the three packaged lactating-cow demonstrations in the public University of Guelph `nasem-dairy` 1.0.2 implementation [46]. Tolerances and mappings were fixed before comparison. An independent, hand-written implementation also checked the selected chain (Supplementary Methods S29).

Shortage magnitude was analysed post hoc using the second extension's saved states and all 96 full-pool and measurement policy–cell records, without changing policies or targets. For canonical signed margin m_j, shortage was $\max(0, -m_j)$. The verdict tolerance classified violations but was not subtracted from shortage magnitude. States with unknown verdicts had undefined shortage magnitudes and remained only

in the failure counts. NEL and CP shortages were normalised by their targets, Ca/P by each state's factorial requirement, and Table 5-1 constraints by their absolute bound. We report mean shortage over all states with defined magnitudes, together with P95 over those states. The top-5% mean (empirical CVaR95) used in the earlier analyses equals this mean times $n_{\text{defined}}/\lceil 0.05 n_{\text{defined}} \rceil$ whenever fewer than 5% of defined states have a positive shortage. This condition held in every full-pool and measurement record at the original coefficient (Supplementary S35).

The same analysis included historical M0 rations, which were distinct from the earlier fixed rule selected by training reference risk. A post-hoc sensitivity analysis held rations fixed and changed only the selected energy chain's microbial-protein coefficient fMCP by ±10% and ±20%, with $\partial \text{NEL}/\partial f_{\text{MCP}} = -0.00218724D$ (Mcal/day per g/kg of DM intake). The five multipliers $(0.8, 0.9, 1.0, 1.1, 1.2)$ were then applied to all 96 full-pool and measurement records, shifting the energy margin by $-0.00218724D(f - f_0)$. Bounds were calculated in a separate diagnostic family (Supplementary S35). Assigned rations in all 78 later-protocol cells were also re-evaluated with the original checker in a separate environment.

4.11 Use of generative artificial intelligence

Generative AI tools (Anthropic Claude Code, OpenAI Codex and OpenAI ChatGPT) were used for English-language editing and to assist with preparing and formatting tables and figures. The authors reviewed the revised text and checked the tables and figures against the underlying study materials. The authors take full responsibility for the accuracy and integrity of the final manuscript.

4.12 Ethics declaration

This simulation study involved no experiments on live animals, human participants or individual farm records.

Data availability

Aggregate tables supporting all numerical results are provided as Supplementary Data and at https://github.com/eeliwuei/dairy-ration-reliability. They include failure counts and rates, paired comparisons, finite-pool lower bounds, descriptive shortage summaries, fixed-action coefficient sensitivity and the LP re-solve control. The same sources provide the as-fed rations for all fixed rules (36 vectors, rounded to one decimal place) and the requirement amounts calculated from the NASEM equations for the reference cow (Supplementary S38).

The composition means and standard deviations transcribed from NASEM (2021), the printed CP target and Table 5-1 bounds, fitted composition models, per-state choices of the nutrient-informed rules and saved simulation states are not redistributed because the source tables are not ours to republish. Under its stated interval assumptions, the identifiability analysis in Supplementary S38 did not uniquely determine any ingredient DM fraction from the released ration vectors. Restricted inputs are retained in the study archive and are available from the corresponding author for editorial

and peer review, subject to the licence conditions of the source tables. Supplementary Methods S30 describes each material class and the checks it supports.

Code availability

The formulation and evaluation engine, tests, specification files with restricted values withheld, frozen protocol digests for the later experiments, statistical verification and saved-state analysis scripts, observed-composition LP re-solve control, and scripts for plotting the aggregate results are available at https://github.com/eeliwuei/dairy-ration-reliability. The repository reproduces the statistical checks and figures from the aggregate tables. Regenerating the simulations also requires the restricted inputs described above.

Acknowledgements

This work was supported by the TIPS Start-up Growth Technology Development (R&D) programme (grant RS-2023-00303920) and the TIPS Overseas Marketing programme (grant 20404166). The funders had no role in study design, analysis, interpretation or the decision to publish.

Author contributions

W.L. conceived the study, developed the methods and software, performed the simulations and analyses, and wrote the first draft. S.J. supervised the work. Both authors interpreted the results, revised the manuscript and approved the submitted version.

Competing interests

Both authors are employees of DeepFarm Inc., a company that develops precision-livestock software. The authors declare no other competing interests.

References

[1] Black, J. R. & Hlubik, J. Basics of computerized linear programs for ration formulation. *Journal of Dairy Science* **63**, 1366–1378 (1980). URL http://dx.doi.org/10.3168/jds.S0022-0302(80)83090-6.

[2] VandeHaar, M. J. & Black, J. Ration formulation using linear programming. *Veterinary Clinics of North America: Food Animal Practice* **7**, 541–556 (1991). URL http://dx.doi.org/10.1016/S0749-0720(15)30788-X.

[3] Kertz, A. Variability in delivery of nutrients to lactating dairy cows. *Journal of Dairy Science* **81**, 3075–3084 (1998). URL http://dx.doi.org/10.3168/jds.S0022-0302(98)75872-2.

[4] Hristov, A., Mertens, D., Zaman, S., Vander Pol, M. & Price, W. Variability in feed and total mixed ration neutral detergent fiber and crude protein analyses

among commercial laboratories. *Journal of Dairy Science* **93**, 5348–5362 (2010). URL http://dx.doi.org/10.3168/jds.2010-3333.

[5] St-Pierre, N. & Weiss, W. Partitioning variation in nutrient composition data of common feeds and mixed diets on commercial dairy farms. *Journal of Dairy Science* **98**, 5004–5015 (2015). URL http://dx.doi.org/10.3168/jds.2015-9431.

[6] Turiello, P., Larriestra, A., Bargo, F., Relling, A. & Weiss, W. Sources of variation in corn silage and total mixed rations of commercial dairy farms. *The Professional Animal Scientist* **34**, 148–155 (2018). URL http://dx.doi.org/10.15232/pas.2017-01704.

[7] Cherney, D., Digman, M. & Cherney, J. Day-to-day variation in forage and mixed diets in commercial dairy farms in New York. *Applied Animal Science* **37**, 11–20 (2021). URL http://dx.doi.org/10.15232/aas.2020-02105.

[8] Barrientos-Blanco, J. A. *et al.* Partitioning of nutrient variation in alfalfa and corn silage by source on New York dairy farms. *Journal of Dairy Science* **107**, 5722–5737 (2024). URL http://dx.doi.org/10.3168/jds.2023-24287.

[9] Duplessis, M., Dubois, K. & Royer, I. Day-to-day variation in trace and macromineral concentrations in corn and mixed grass–legume silages of Canadian commercial dairy herds. *Canadian Journal of Animal Science* **104**, 356–366 (2024). URL http://dx.doi.org/10.1139/cjas-2023-0121.

[10] Yoder, P., St-Pierre, N. & Weiss, W. A statistical filtering procedure to improve the accuracy of estimating population parameters in feed composition databases. *Journal of Dairy Science* **97**, 5645–5656 (2014). URL http://dx.doi.org/10.3168/jds.2013-7724.

[11] Tran, H. *et al.* Development of feed composition tables using a statistical screening procedure. *Journal of Dairy Science* **103**, 3786–3803 (2020). URL http://dx.doi.org/10.3168/jds.2019-16702.

[12] Schlageter-Tello, A. *et al.* ASAS-NANP symposium: Ruminant/nonruminant feed composition: Challenges and opportunities associated with creating large feed composition tables. *Journal of Animal Science* **98** (2020). URL http://dx.doi.org/10.1093/jas/skaa240.

[13] Charnes, A. & Cooper, W. W. Chance-constrained programming. *Management Science* **6**, 73–79 (1959). URL http://dx.doi.org/10.1287/mnsc.6.1.73.

[14] St-Pierre, N. & Harvey, W. Incorporation of uncertainty in composition of feeds into least-cost ration models. 1. single-chance constrained programming. *Journal of Dairy Science* **69**, 3051–3062 (1986). URL http://dx.doi.org/10.3168/jds.S0022-0302(86)80768-8.

[15] St-Pierre, N. & Harvey, W. Incorporation of uncertainty in composition of feeds into least-cost ration models. 2. joint-chance constrained programming. *Journal of Dairy Science* **69**, 3063–3073 (1986). URL http://dx.doi.org/10.3168/jds.

s0022-0302(86)80769-x.

[16] Roush, W. B., Stock, R. H., Cravener, T. L. & D'Alfonso, T. H. Using chance-constrained programming for animal feed formulation at Agway. *Interfaces* **24**, 53–58 (1994). URL http://dx.doi.org/10.1287/inte.24.2.53.

[17] Tozer, P. Least-cost ration formulations for Holstein dairy heifers by using linear and stochastic programming. *Journal of Dairy Science* **83**, 443–451 (2000). URL http://dx.doi.org/10.3168/jds.S0022-0302(00)74901-0.

[18] Weiss, W. & St-Pierre, N. Perspective and Commentary: Variation in nutrient composition of feeds and diets and how it can affect formulation of dairy cow diets. *Applied Animal Science* **40**, 608–618 (2024). URL http://dx.doi.org/10.15232/aas.2024-02578.

[19] Soyster, A. L. Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* **21**, 1154–1157 (1973). URL http://dx.doi.org/10.1287/opre.21.5.1154.

[20] Ben-Tal, A. & Nemirovski, A. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* **88**, 411–424 (2000). URL http://dx.doi.org/10.1007/PL00011380.

[21] Bertsimas, D. & Sim, M. The price of robustness. *Operations Research* **52**, 35–53 (2004). URL http://dx.doi.org/10.1287/opre.1030.0065.

[22] Tozer, P. & Stokes, J. A multi-objective programming approach to feed ration balancing and nutrient management. *Agricultural Systems* **67**, 201–215 (2001). URL http://dx.doi.org/10.1016/S0308-521X(00)00056-1.

[23] Uyeh, D. D. *et al.* Precision animal feed formulation: An evolutionary multi-objective approach. *Animal Feed Science and Technology* **256**, 114211 (2019). URL http://dx.doi.org/10.1016/j.anifeedsci.2019.114211.

[24] Li, J. *et al.* The application of nonlinear programming on ration formulation for dairy cattle. *Journal of Dairy Science* **105**, 2180–2189 (2022). URL http://dx.doi.org/10.3168/jds.2021-20817.

[25] Usigbe, M. J., Uyeh, D. D., Park, T., Ha, Y. & Mallipeddi, R. Many objective optimization and decision support for dairy cattle feed formulation. *Scientific Reports* **15** (2025). URL http://dx.doi.org/10.1038/s41598-025-96633-z.

[26] Nemhauser, G. L., Wolsey, L. A. & Fisher, M. L. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming* **14**, 265–294 (1978). URL https://doi.org/10.1007/BF01588971.

[27] Mandi, J., Bucarey, V., Tchomba, M. M. K. & Guns, T. Decision-focused learning: Through the lens of learning to rank. *Proceedings of Machine Learning Research (ICML 2022)* **162**, 14935–14947 (2022). URL https://proceedings.mlr.press/v162/mandi22a.html.

[28] de Souza, R., Tempelman, R., Allen, M. & VandeHaar, M. Updating predictions of dry matter intake of lactating dairy cows. *Journal of Dairy Science* **102**, 7948–7960 (2019). URL http://dx.doi.org/10.3168/jds.2018-16176.

[29] NASEM. *Nutrient Requirements of Dairy Cattle: Eighth Revised Edition* (National Academies Press, 2021). URL http://dx.doi.org/10.17226/25806. National Academies of Sciences, Engineering, and Medicine (NASEM).

[30] USDA Agricultural Marketing Service. Livestock, Poultry and Grain Market News: National Grain and Oilseed Processor Feedstuff Report (AMS 3511, report for 14–18 September 2026); Iowa Daily Cash Grain Bids (AMS 2850, 23 September 2026); Missouri Direct Hay Report (AMS 2929, week ending 18 September 2026). https://mymarketnews.ams.usda.gov (2026). URL https://mymarketnews.ams.usda.gov. Reports AMS 3511, 2850 and 2929; accessed 25 September 2026.

[31] University of Missouri Extension. FEEDVAL II: by-product feed values for lactating dairy cows (page dated 17 September 2026). http://agebb.missouri.edu/dairy/byprod/lactation.php (2026). URL http://agebb.missouri.edu/dairy/byprod/lactation.php. Accessed 25 September 2026.

[32] USDA Agricultural Marketing Service. Organic Cotton Market Summary, volume 17 (August 2026): conventional and organic cottonseed price ranges, 2025 crop. https://www.ams.usda.gov/mnreports/cnaocms.pdf (2026). URL https://www.ams.usda.gov/mnreports/cnaocms.pdf. Cotton and Tobacco Program, Cotton Market News Division, Memphis, TN; accessed 28 September 2026.

[33] Edwards, W. & Hart, C. E. Pricing forage in the field (Ag Decision Maker file a1-65). https://www.extension.iastate.edu/agdm/crops/html/a1-65.html (2018). URL https://www.extension.iastate.edu/agdm/crops/html/a1-65.html. Updated September 2018; accessed 25 September 2026.

[34] Beuerlein, J. & Lindsey, A. J. Bushels, test weights, and calculations (AGF-503). https://cfaes.osu.edu/fact-sheet/bushels-test-weights-and-calculations (2018). URL https://cfaes.osu.edu/fact-sheet/bushels-test-weights-and-calculations. Revised 26 December 2018; accessed 25 September 2026.

[35] Franco-Hernández, M. A., Corona, L., Plascencia, A. & Zinn, R. A. Practical parameters for assessing starch digestion and feeding value of steam-flaked corn in finishing diets for feedlot cattle. *Journal of Applied Animal Research* **51**, 24–30 (2023). URL http://dx.doi.org/10.1080/09712119.2022.2149537.

[36] Plascencia, A. *et al.* Influence of cracked, coarse ground or fine ground corn on digestion, dry matter intake and milk yield in Holstein cows. *Journal of Applied Animal Research* **35**, 149–154 (2009). URL http://dx.doi.org/10.1080/09712119.2009.9707005.

[37] Luedtke, J. & Ahmed, S. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization* **19**, 674–699 (2008).

URL http://dx.doi.org/10.1137/070702928.

[38] Pagnoncelli, B. K., Ahmed, S. & Shapiro, A. Sample Average Approximation method for chance constrained programming: Theory and applications. *Journal of Optimization Theory and Applications* **142**, 399–416 (2009). URL http://dx.doi.org/10.1007/s10957-009-9523-6.

[39] Huangfu, Q. & Hall, J. A. J. Parallelizing the dual revised simplex method. *Mathematical Programming Computation* **10**, 119–142 (2018). URL http://dx.doi.org/10.1007/s12532-017-0130-5.

[40] Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17**, 261–272 (2020). URL http://dx.doi.org/10.1038/s41592-019-0686-2.

[41] He, Z. On the error rate of conditional Quasi–Monte Carlo for discontinuous functions. *SIAM Journal on Numerical Analysis* **57**, 854–874 (2019). URL https://doi.org/10.1137/18M118270X.

[42] Liu, S. Conditional Quasi-Monte Carlo with constrained active subspaces. *SIAM Journal on Scientific Computing* **46**, A2999–A3021 (2024). URL https://doi.org/10.1137/23M1548918.

[43] Gilbert, A. D., Kuo, F. Y. & Sloan, I. H. Preintegration is not smoothing when monotonicity fails. *Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer (Botev, Z., Keller, A., Lemieux, C. and Tuffin, B., eds; Springer, Cham)* 169–191 (2022). URL https://doi.org/10.1007/978-3-031-10193-9_9.

[44] Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413 (1934). URL http://dx.doi.org/10.1093/biomet/26.4.404.

[45] Carroll, R. J., Ruppert, D., Stefanski, L. A. & Crainiceanu, C. M. *Measurement Error in Nonlinear Models: A Modern Perspective* 2 edn (Chapman and Hall/CRC, Boca Raton, 2006). URL http://dx.doi.org/10.1201/9781420010138.

[46] Centre for Nutrition Modelling, University of Guelph. NASEM Model Python, version 1.0.2. https://github.com/CNM-University-of-Guelph/NASEM-Model-Python/releases/tag/v1.0.2 (n.d.). Pinned software release; accessed 2 October 2026.

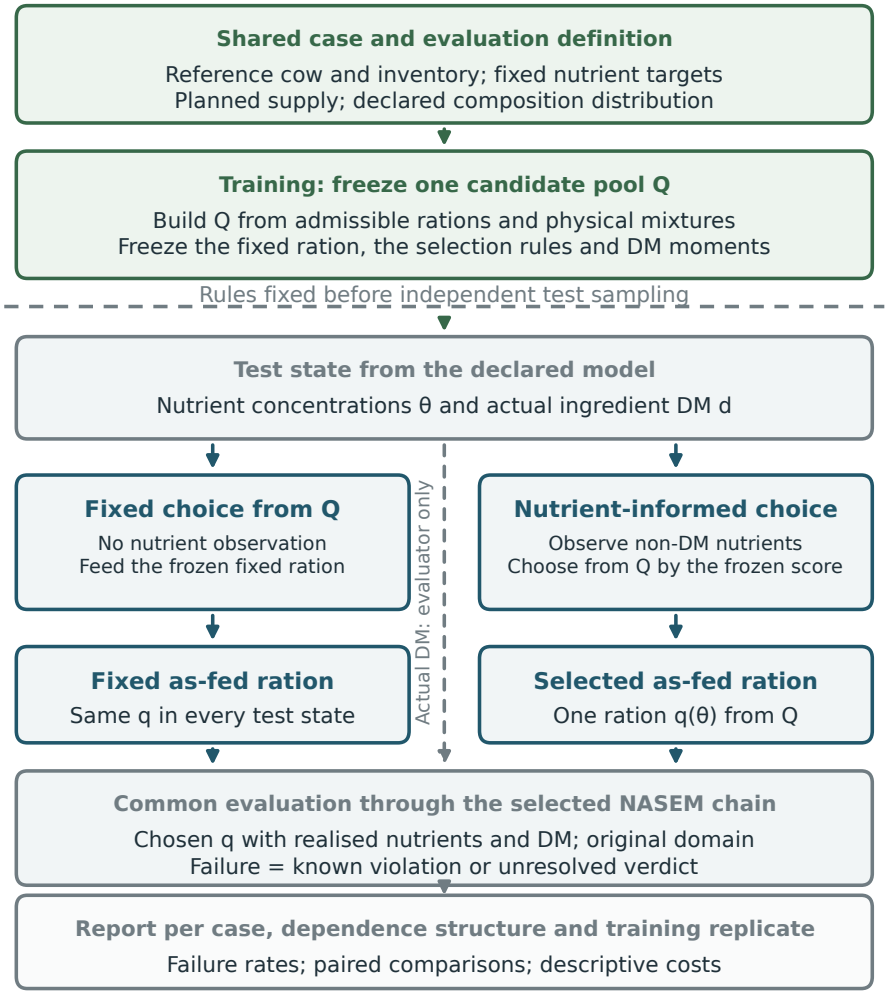


Fig. 1. Decision structure and information timing for fixed and nutrient-informed ration selection. The candidate pool, the fixed rule's single ration and the nutrient-informed rule are frozen after training and before independent testing. At test time, the nutrient-informed rule observes non-DM nutrient concentrations before choosing a ration. Actual dry matter (DM) enters only when the chosen ration is evaluated. Both rules are evaluated against the same nutritional event under the selected model; a known violation or an unresolved verdict counts as failure. DM-only information experiments use a separate observation interface (Methods).

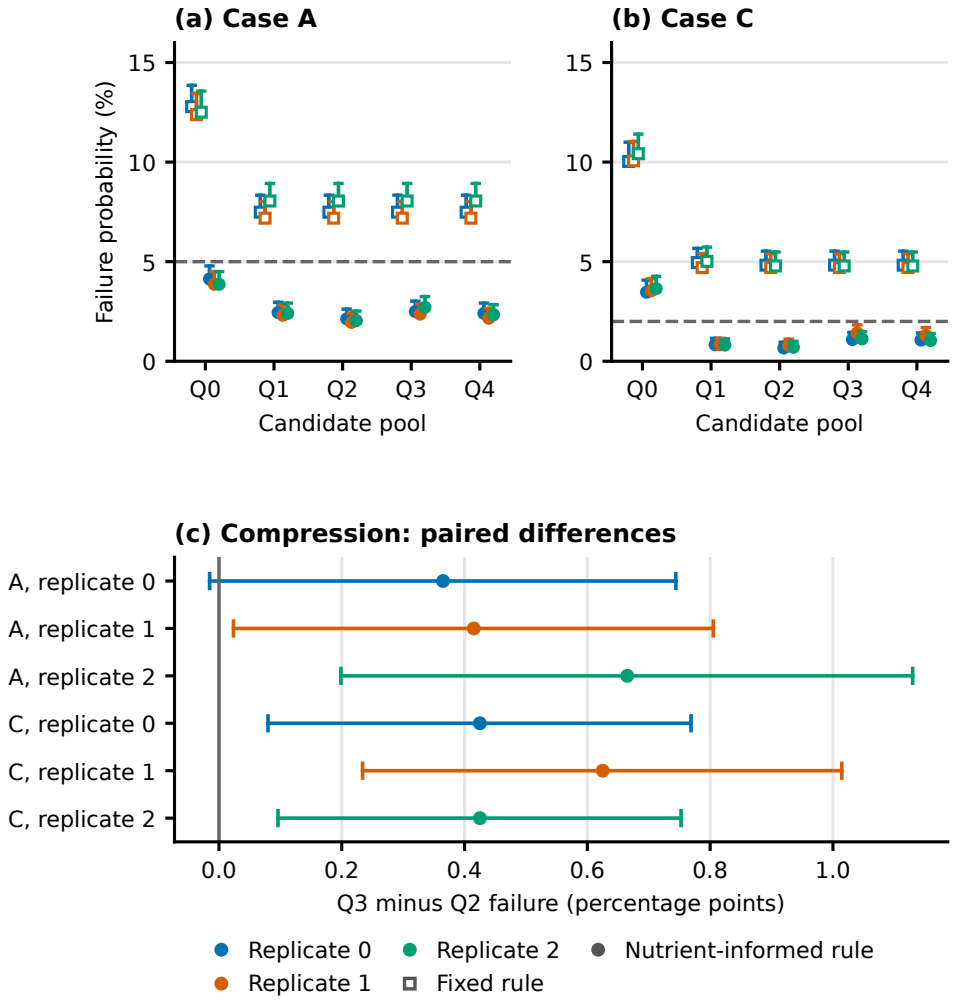


Fig. 2. Fixed and nutrient-informed selection on five matched candidate pools under TAB variability and dependence structure C0. Q_0 is the original admissible pool; Q_1 adds all admissible earlier price candidates; Q_2 adds training-defined physical mixtures; Q_3 compresses Q_2 by success coverage; and Q_4 restores all Q_0 rations to Q_3 without a size cap. Within each cell, the fixed and nutrient-informed (delta) rules share the original reference event, animal targets, planned supply and test states. Points and one-sided upper confidence limits are shown separately for each of the three trained replicates. Panel (c) shows paired failure differences for Q_3 minus Q_2; an interval above zero indicates higher risk after compression. Five of the six A/C compression comparisons had adjusted paired intervals entirely above zero. Rate and paired inference families have separate error budgets and do not support a combined 95% claim.

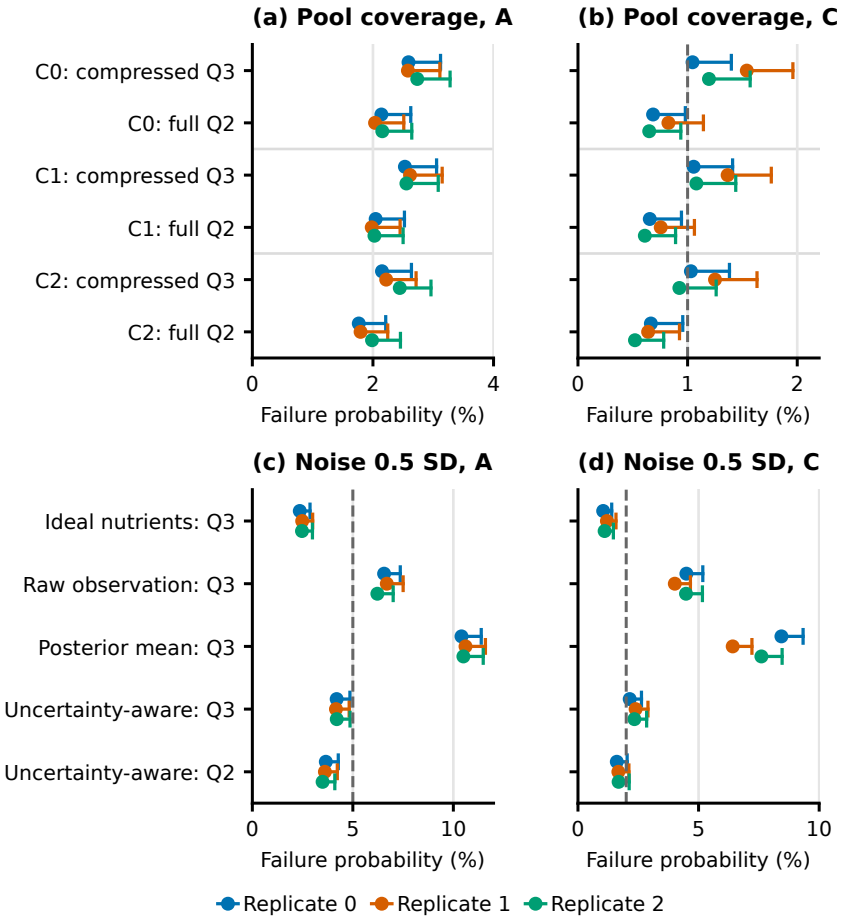


Fig. 3. Failure risk in the candidate-coverage (a,b) and measurement-uncertainty (c,d) experiments, which used separate protocols. Panels (a,b) compare nutrient-informed selection on the full Q_2 and compressed Q_3 pools in fresh C0/C1/C2 tests; all 18 adjusted pairs favoured full coverage. Panels (c,d) use different C0 tests with hypothetical noise at 0.5 training SD, except for the ideal control. On Q_3, using posterior means alone increased failure relative to raw observations, whereas adding posterior variance to the score (the uncertainty-aware rule) improved all six paired comparisons. The full Q_2 versus Q_3 comparison in panels (c,d) uses the same uncertainty-aware score; there is no raw-observation control on Q_2. Points and endpoints show risks and one-sided binomial upper confidence bounds for each training replicate ($n = 20,000$). Bounds are adjusted within their respective inference families, and replicates are not pooled. Dashed lines mark C's 1% target in (b), A's 5% in (c) and C's 2% in (d). A's coverage risks lie below its 5% target, outside the plotted scale. Both A and C meet the adjusted 5% criterion, but C's high-noise upper confidence bounds exceed its stricter 2% criterion.

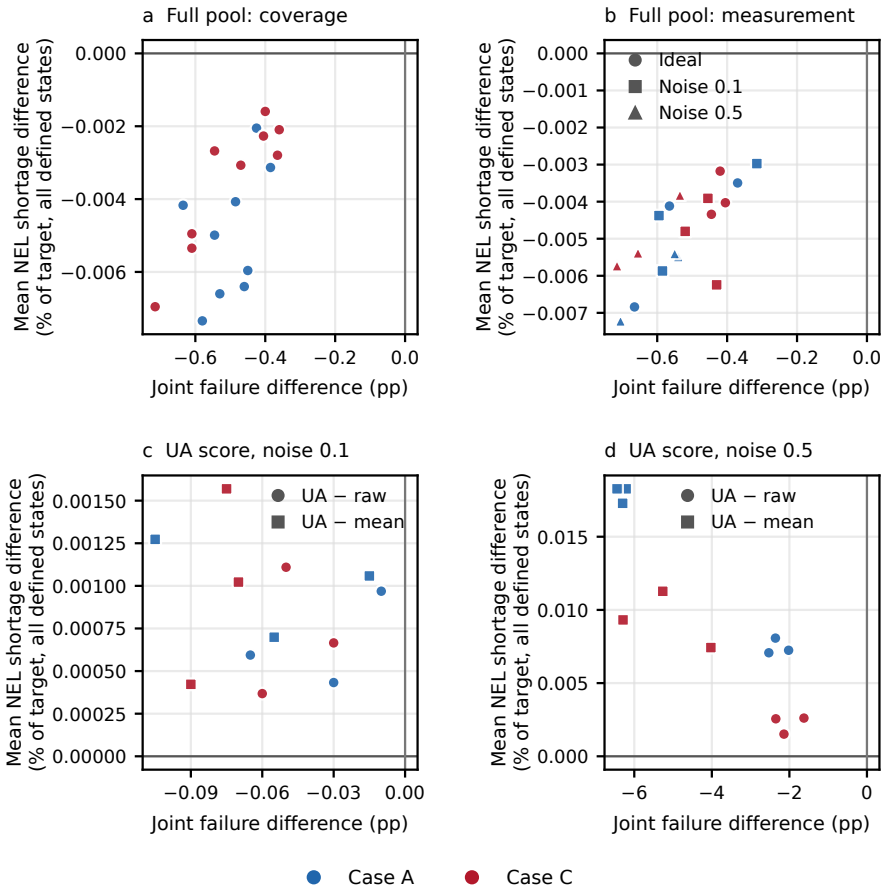


Fig. 4. Paired differences between selection rules in failure frequency under the model and mean NEL shortage; improvement in one does not ensure improvement in the other. Each point compares two rules within one protocol. Axes show policy minus comparator: the observed failure difference (percentage points) and the difference in mean NEL shortage over all states with defined magnitudes (percent of the original target). Panel (a) shows full Q_2 minus compressed Q_3 in all 18 coverage cells, and panel (b) the same contrast under ideal, 0.1-SD and 0.5-SD observation conditions (18 points). Panels (c,d) show uncertainty-aware (UA) minus raw-observation or posterior-mean selection on Q_3 at the two noise levels (12 points each); in panel (c), the adjusted failure-difference intervals included zero. At high noise, UA selection on Q_3 increased NEL violation incidence relative to raw selection but lowered the mean and P95 shortage conditional on violation; changes in maxima varied in direction. Colours identify cases, and scales differ among panels. Shortage differences are descriptive and were not subjected to new significance tests. Results for all nine constraints, including denominators and counts of states with undefined magnitudes, are reported in Supplementary Methods S35.

Tables

Table 1: Definitions of cases R0, A, B and C and the scope of the model constraints. Exact animal-specific inputs are provided in the restricted reproducibility material.

Item	Definition	Scope or interpretation
R0	Original eight-ingredient inventory	Planned supply at NASEM Eq. 2-1 animal-factor prediction
A	R0 plus 10% planned DM	Same NEL, CP, Ca/P and carbohydrate targets; all supplied DM assumed consumed
B	R0 plus whole cottonseed	Nine ingredients; cottonseed capped at 10% planned DM as a project assumption
C	A and B combined	No further planned-DM increase in subsequent candidate-pool experiments
NEL	Selected Chapter 3 chain	Fixed fMCP approximation; not a complete microbial/protein model
CP	Table 21-3 illustrative intake	Case target, not a validated MP/EAA/RDP requirement
Ca/P	Absorbed supply and factorial needs	Fixed absorption inputs; maintenance needs use realised supplied DM
NDF/starch	Five Table 5-1 rules	Conditional recommendations with a study-set planned corn-starch premise
Structure	Supply equality and planning caps	Retained outside the failure budget; planning CP upper/EE limits are not main-event members

Table 2: Experimental protocols, candidate pools, observations, scores, and training and test populations. Each test cell in the pool-based protocols used 20,000 states per independently trained replicate; replicates were analysed separately. The exploratory LP control used three development streams per case and did not train a policy on a finite pool. Comparisons were paired only within the same protocol and target population. Each inference family (rate, paired, pool-bound and exploratory) had a separate error allocation.

Experiment	Frozen candidate pool	Observation and score	Training and test populations
Attribution (S31)	Q_0–Q_4; one fixed rule per pool	Fixed, or ideal nutrients with delta	C0; TAB and NARROW; A/C plus R0/B controls
Full vs compressed (S35)	Attribution Q_2 versus Q_3	Ideal nutrients; unchanged delta	C0 training; C0/C1/C2 tests
Measurement (S35)	Same C0 parent Q_2/Q_3	Ideal, raw-observation, posterior-mean or uncertainty-aware; 0.1 or 0.5 SD	C0 training and tests; A/C
Forward transfer (S33)	Frozen C0 compressed pool	None, DM-only or nutrients; 64-node plug-in; delta controls	C0 training; C0/C1/C2 tests
Reverse transfer (S33)	Historical C2 expanded pool	Fixed or ideal nutrient delta	C2 training; C2/C0 tests
Generated rations (S36)	Attribution-stage Q_3 plus new training-generated candidate rations	Fixed or ideal nutrient delta	C0 training and tests; R0/B
LP re-solve control (S37)	None: continuous re-formulation at the observed composition	Ideal nutrients; least-cost LP, without or with a DM margin	A/C; TAB/C0; three development streams (exploratory)

Actual DM was hidden from nutrient-informed selection; DM-only rules used a separate observation interface. The identity of the S36 parent pools was verified by comparing ration vectors, as documented in its lineage table.

Table 3: Failure risks and feed costs relative to M0 for fixed and nutrient-informed (delta) selection in the matched-pool experiment under TAB variability and dependence structure C0. Results are shown for three separately trained replicates. Within each cell, both rules use the same candidate pool and test states.

Case	Pool	Candidates	Fixed max. upper (%)	Delta risk range (%)	Delta max. upper (%)	Fixed cost ratio to M0	Delta cost ratio to M0
A	Q_0	332	13.8553	3.865–4.130	4.7850	1.169–1.180	1.133
A	Q_1	515	8.9251	2.310–2.445	2.9622	1.117–1.151	1.112–1.113
A	Q_2	599	8.9251	1.950–2.130	2.6161	1.117–1.151	1.113–1.115
A	Q_3	18–24	8.9251	2.365–2.705	3.2463	1.117–1.151	1.112–1.119
A	Q_4	347–349	8.9251	2.160–2.405	2.9184	1.117–1.151	1.114–1.119
C	Q_0	353	11.4032	3.465–3.645	4.2642	1.228	1.147
C	Q_1	894	5.7295	0.815–0.830	1.1500	1.099–1.108	1.126–1.127
C	Q_2	978	5.5273	0.665–0.795	1.1090	1.134–1.141	1.130–1.134
C	Q_3	11–15	5.5273	1.090–1.420	1.8254	1.134–1.141	1.124–1.135
C	Q_4	363–367	5.5273	1.040–1.300	1.6898	1.134–1.141	1.123–1.135

Q_0: original admissible pool; Q_1: plus admissible earlier price candidates; Q_2: plus physical mixtures; Q_3: after coverage compression; Q_4: uncapped $Q_0 \cup Q_3$. Ranges and maximum upper confidence bounds summarise the 20,000-state replicates without pooling them. Cost ratios divide the mean as-fed cost of the rations chosen by each rule by the cost of that case's nominal least-cost ration M0. This is also the baseline for the LP re-solve control (ranges across replicates; the Q_0 fixed ration costs 1.169–1.180 times M0 in A and 1.228 in C). Every Q_3 replicate meets the criterion at 5% in A and 2% in C. In these matched C0 attribution cells, the uncompressed C pool (Q_2) meets the 1% criterion in only two replicates (upper confidence bounds 0.955724%, 1.109024% and 0.997209%). Attainment of these criteria was also checked using integer binomial tails. Replicates are not pooled, and the confidence families are not combined.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData.zip
- SupplementaryInformation.pdf