agronomy

MDPI

Article

Stage-Dependent Heritability and Multivariate Organization of Physiological and Canopy Phenomic Traits in Tropical Soybean

José Ignacio Rodríguez-Valencia [1,*], Gustavo Adolfo Marín-Ramírez [2,3], Diego Felipe Conejo Rodríguez [1], Jhony Armando Benavides-Bolaños [4], Edwin Erazo-Mesa [4] and Mario Augusto García Dávila [1]

1 Faculty of Agricultural Sciences, Universidad Nacional de Colombia, Palmira Campus, Palmira 763533, Colombia; dfconejoro@unal.edu.co (D.F.C.R.)
2 Centro Nacional de Investigaciones de Café (Cenicafé), Federación Nacional de Cafeteros de Colombia, Chinchiná 170009, Colombia; gustavo.marin@ufv.br
3 Laboratory of Plant-Pathogen Genomics and Genetics, Department of Plant Pathology, BIOAGRO, Federal University of Viçosa, Viçosa–Minas Gerais 36570-900, Brazil
4 Escuela EIDENAR, Facultad de Ingeniería, Universidad del ValleCalle 13 # 100-00, Edificio E39, Cali 760032, Colombia; jhony.benavides@correounivalle.edu.co (J.A.B.-B.); osvaldo.erazo@correounivalle.edu.co (E.E.-M.)
* Correspondence: jirodriguezv@unal.edu.co

Abstract

Phenomic traits can support soybean breeding when genetic signal, temporal redundancy, and uncertainty are evaluated together. We assessed 30 tropical soybean genotypes in ELO_25A using chlorophyll fluorescence and SPAD measurements, UAV-derived RGB indices, and agronomic traits across developmental stages. Spatial mixed models, with PAR as the sole physiological covariate, provided adjusted means, BLUPs, Cullis generalized heritability, and prediction-error covariance. Targeted reduced-rank models assessed temporal structure, and equal-weight PCA with propagated uncertainty summarized six retained features. Genetic precision varied with development and was strongest at R3–R4. Model comparisons favored rank-one summaries for SPAD V5–R3 and GLI R3–R4–R6, with proportionality imposed by the selected structure, whereas SPAD R6 provided distinct late information. PC1 explained 39.12% of standardized variation and contrasted R3 energy regulation with chlorophyll status. It was associated with plant height and pod number ($r = 0.550$ and 0.560; both $q = 0.006$) but not yield ($r = 0.056$; $q = 0.822$). Genotype scores formed a continuum, and no cluster solution met the stability criteria. The results support stage-aware feature reduction, germplasm discrimination, and efficient phenotyping within ELO_25A, but do not validate yield prediction or indirect selection. Validation across environments and independent populations is required before use in plant breeding.

Keywords: plant breeding; chlorophyll fluorescence; Cullis heritability; genetic covariance; RGB vegetation indices; soybean

Check for updates

Academic Editor: Junhua Peng

Received: 29 September 2026
Revised: 27 October 2026
Accepted: 30 October 2026
Published: 12 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

High-throughput phenotyping (HTP) allows soybean breeders to observe plant function and canopy development before yield and other end-of-season traits are available. Its value depends on whether those measurements are biologically interpretable and practical to collect at breeding scale [1–3]. MultispeQ records describe photosystem II regulation and energy dissipation [4–6], while UAV-RGB imagery captures plot-level differences in canopy greenness and development [7,8]. Soybean applications have been encouraging, but most were developed with large populations, repeated environments, spatially adjusted

trials, or canopy time series [9–13]. Evidence obtained under those conditions cannot be assumed to apply to a small tropical panel grown in one environment.

For breeding decisions, repeatability is only one part of the evidence. A trait may differ among genotypes and still duplicate another measurement, appear during a narrow developmental window, or have little connection with the target outcome. Conversely, lack of association with yield does not remove its physiological meaning. The trait may still distinguish lines, reveal a response not captured by other sensors, or help determine when field measurements are worth taking. In an early-stage program, a small set of well-timed measurements may therefore be more useful than a high-dimensional prediction system [14,15].

Timing complicates this assessment. Stagewise heritability measures the precision of genetic differentiation at one stage but says nothing about whether the same genotypic pattern persists at another. Temporal covariance models address that gap by distinguishing shared expression from stage-specific behavior. This distinction is useful when the aim is to remove redundant measurements or identify an earlier proxy. Reduced-rank structures are especially attractive for small panels because they require fewer covariance parameters, although their rank constraint limits the interpretation of the fitted correlations.

Tropical soybean breeding in Colombia offers a relevant setting for this evaluation. Regional work has documented agronomic variation and the usefulness of mixed-model prediction in locally adapted germplasm [16–18]. Programs working under these conditions also need affordable ways to combine field physiology, consumer-grade RGB UAVs, and conventional agronomic records. In a single-environment experiment, however, the analysis can describe genetic organization within that trial but cannot demonstrate multi-environment selection value.

The immediate questions are whether genotypes can be distinguished with adequate precision, which stages carry repeated or complementary information, and whether the resulting phenomic pattern agrees with agronomic performance in the same environment.

We addressed these questions by estimating stage-specific genetic signal after spatial adjustment, then modeling temporal covariance for the repeated traits that remained estimable. A reduced phenomic panel was assembled from nonredundant physiological and UAV-RGB measurements, and uncertainty in the adjusted means was carried into the multivariate analysis. Finally, we examined concordance with yield, plant height, and pod number within ELO_25A. We expected some measurements to describe the same temporal dimension and others to add stage-specific information; neither outcome was assumed to imply yield association or discrete genotype groups.

2. Materials and Methods

2.1. Field Environment, Plant Material and Design

We conducted the ELO_25A trial during the first semester of 2025 at La Eloísa Experimental Station in Obando, Valle del Cauca, Colombia (4.6202° N, 75.9955° W; 917 m a.s.l.). The crop was sown on 15 April and harvested on 12 August 2025, spanning a 119-day field cycle; see Figure 1. Daily weather records from NASA POWER were used only to characterize the growing period and were not interpreted as plot-level microclimatic observations [19].

The panel comprised 30 soybean (*Glycine max* [L.] Merr.) genotypes: 25 proprietary experimental lines from the GS series and five reference materials. The lines were supplied by Granbras Sementes Ltda. (commercially known as Great Seed; Paracatu, Minas Gerais, Brazil). According to information provided by the company, they were selected from an initial set of 100 experimental materials based on their agronomic adaptation and favorable grain-yield performance in evaluations conducted in Brazil. ELO_25A represented their

first evaluation under Colombian growing conditions, with the purpose of characterizing their agronomic and phenomic performance and examining their adaptation to the target environments. Thus, the panel was assembled for breeding evaluation rather than to represent or maximize soybean genetic diversity.

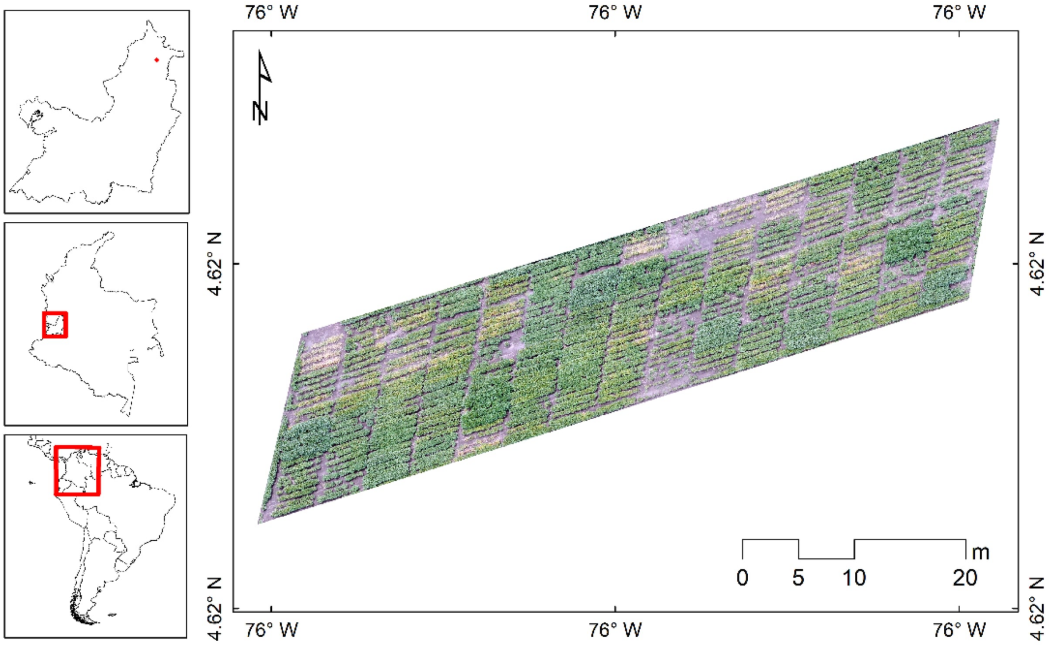


Figure 1. Geographic setting and UAV orthomosaic of the ELO_25A soybean experiment in Colombia. The left panels locate the experimental site within South America, Colombia, and Valle del Cauca; the red point and red square identify the study location at the corresponding geographic scales. The main panel shows the RGB orthomosaic produced from UAV imagery acquired during the 2025 growing season.

The reference materials included three commercial cultivars, identified in the experimental records and analyses as NS7923, NS6323, and FNS-01, and two farmer-maintained materials, A_1 and Monsoy. The latter had been grown locally for several years and were included as local references based on their favorable agronomic performance, as reported by farmers. Their field identifiers were retained because their formal cultivar denominations could not be verified. The parental pedigrees of the GS lines were not disclosed by the germplasm provider because of commercial confidentiality restrictions; consequently, pedigree-based relationships among the materials in the panel were not evaluated.

We arranged the trial as a randomized alpha-lattice with three replicates, six incomplete blocks within each replicate, and five genotypes per incomplete block. Each of the 90 plots consisted of four 5 m rows spaced 0.60 m apart. Physiological measurements came from the two central rows, whereas UAV and agronomic variables were summarized at plot level. Inference was restricted to the evaluated genotype panel at this location and during the 2025 growing season. The stagewise heritability estimates describe the precision of genotype differentiation within ELO_25A, while the temporal models summarize relationships among measurements from that same experiment. With no replication across locations or seasons, the design could not separate genotype effects that persist across environments from genotype-specific responses to the conditions of ELO_25A, or establish whether the temporal relationships would recur elsewhere.

2.2. Stage-Resolved Phenotyping

Physiological measurements and UAV acquisitions were conducted on common dates across plots. Developmental stages were visually assessed using the method outlined in [20] and recorded in the fieldbook together with sampling dates and days after sowing (DAS). For R1, more than 50% of the plants in each plot were confirmed to have begun flowering.

Physiological measurements were collected on 21 May at V5 (36 DAS), 4 June at R1 (50 DAS), 17 June at R3 (63 DAS), and 9 July 2025 at R6 (85 DAS). UAV acquisitions occurred on 20 May at V5 (36 DAS), 4 June at R1 (50 DAS), 19 June at R3 (63 DAS), 25 June at R4 (71 DAS), and 9 July at R6 (85 DAS). No physiological measurements were collected at R4.

The recorded stages provided a common developmental reference but did not imply identical dates of stage attainment among genotypes. Residual developmental differences were therefore acknowledged as a potential source of variation in the temporal analyses.

2.2.1. Leaf Physiology

Leaf physiology was measured using a MultispeQ V2.0 instrument (PhotosynQ Inc., East Lansing, MI, USA) [5] at V5, R1, R3, and R6, following the soybean developmental stages described in [20]. Leaf physiology was evaluated using a MultispeQ v2.0 instrument during campaigns nominally assigned to the V5, R1, R3, and R6 stages, according to the developmental scale of [20]. Measurements were performed using the Photosynthesis RIDES 2.0 protocol on the PhotosynQ platform [], as documented in the original data export. No additional protocol identifier was recorded, and the instrument firmware and application versions could not be recovered from the available metadata.

For each plot–stage combination, two plants were assessed, one from each central row (rows 2 and 3). On each plant, the central leaflet of the youngest fully expanded, sun-exposed trifoliolate leaf was measured while remaining attached to the plant. The leaflet was enclosed in the instrument clip. The two observations were treated as subsamples and averaged to obtain a single plot–stage value.

For each plot–stage combination, one plant from each central row (rows 2 and 3) was assessed consecutively on a fully expanded trifoliolate leaf. We treated the two plant observations as subsamples and averaged them to obtain one plot–stage value. Measurements were collected between approximately 10:30 and 14:00 h. This interval is consistent with field applications of MultispeQ measurements during daytime illumination [5] but was not established here as an optimal sampling window. Measurements were scheduled approximately between 10:30 and 14:00 h Colombia Time (COT; UTC−5). This midday-centered window allowed measurements to be conducted under light-adapted daytime conditions while avoiding the comparatively low irradiance of early morning and late afternoon and provided sufficient time to complete approximately 180 observations across the 90 plots. Comparable field studies have conducted light-adapted MultispeQ measurements between 10:00 and 14:00 h or within similar daytime periods [21,22]. However, this interval was not considered physiologically optimal or environmentally constant because the measured fluorescence traits remain responsive to short-term variation in incident radiation. Consequently, PAR was recorded with every observation and included as the only measurement-environment covariate in the physiological models. Restricting collection to this period did not ensure constant irradiance or eliminate other sources of within-day physiological variation.

Measurements followed a serpentine route based on the plot sequence established in the fieldbook and displayed in the Field Book mobile application (PhenoApps), rather than a randomized acquisition sequence [23]. Plots were assessed consecutively within each replicate before proceeding to the next replicate, and the same general traversal pattern was maintained across developmental stages and measurement dates. Within each plot, the

two selected plants, one from each central row (rows 2 and 3), were measured consecutively before proceeding to the next plot.

No formal irradiance thresholds, cloud-related pause rules, or weather-based criteria for repeating measurements were documented. Consequently, transient variation in illumination was not controlled through acquisition pauses or weather-based repetition. PAR was recorded simultaneously with each observation and included as the only observation-level environmental covariate in the physiological models.

We focused on five physiological traits: operating quantum efficiency of photosystem II (ΦPSII), photochemical openness (qL), regulated non-photochemical quenching (NPQt), light-adapted maximum efficiency (Fv′/Fm′), and SPAD chlorophyll index. The SPAD variable corresponded to the relative-chlorophyll output generated directly by the MultispeQ v2.0 instrument during the Photosynthesis RIDES 2.0 measurement and stored in the PhotosynQ export under the field name "SPAD." It was not obtained using a separate chlorophyll meter. The MultispeQ estimates this relative index from leaf absorbance measurements at 650 and 940 nm; therefore, the values were interpreted as relative chlorophyll units rather than absolute chlorophyll concentrations. Because SPAD was recorded within the same instrument record as the fluorescence variables, PAR, plot identifier, stage, and timestamp, no separate matching procedure was required.

2.2.2. UAV-RGB Canopy Phenomics

RGB images were collected at V5, R1, R3, R4, and R6 with a DJI Air 2 UAV flying approximately 30 m above the field. Flights were confined to midday periods and followed a standardized overlap setting. UAV imagery was acquired using a DJI Air 2 equipped with its factory-integrated 12-MP RGB camera during common campaigns at V5 (20 May), R1 (4 June), R3 (19 June), R4 (25 June), and R6 (9 July 2025). Flights began at approximately 12:00 h COT (UTC−5) and lasted approximately 2 – 2.25 min. Missions were flown at a nominal height of 30 m AGL, at 3 m s^{-1}, with 70% forward and 70% side overlap. Source images were acquired in JPG format at 4000 × 3000 pixels using automatic white balance. Exposure mode, shutter speed, and ISO were not recorded in the available image metadata. The achieved ground sampling distances, calculated from the exported GeoTIFF geotransformation metadata, were approximately 9.69, 6.45, 13.79, 13.32, and 9.30 mm $pixel^{-1}$ for V5, R1, R3, R4, and R6, respectively. Plot polygons defined the extraction area, and soil and other background pixels were removed before the indices were calculated. Orthomosaics were reconstructed using Agisoft Metashape Professional v1.8.4 (build 14654). The processing report documented image alignment at medium accuracy, source-based reference preselection, guided matching, adaptive camera-model fitting, a key-point limit of 40,000, and a tie-point limit of 80,000. Depth maps were generated at medium quality with mild filtering, followed by dense-cloud reconstruction and generation of an interpolated digital elevation model. Orthomosaics were generated from this model using mosaic blending and hole filling. Georeferencing relied on the geographic coordinates recorded by the UAV camera. The ground pixel size of the exported orthomosaics ranged from 6.45 to 13.79 mm $pixel^{-1}$.

One polygon was associated with each of the 90 experimental plots for data extraction using the FIELDimageR-QGIS workflow. Pixels corresponding to soil and other background elements were excluded using the threshold-based fieldMask procedure. The arithmetic mean of the retained canopy pixels was subsequently calculated for each combination of plot, developmental stage, and vegetation index. An 18% neutral-gray reference panel was photographed during image acquisition. Vegetation indices were calculated from the digital RGB values of the orthomosaics, and globally min–max-normalized variables

were excluded from the analysis. The same reconstruction, plot delineation, masking, and extraction workflow was applied consistently across the V5, R1, R3, R4, and R6 campaigns.

The UAV analysis used the raw visible-spectrum indices NGRDI [24], VARI [25], GLI [26], ExG [27], and TGI [24]. The expected 450 plot-stage records were complete. Globally min-max-normalized copies were not analyzed. One source field was labeled normalized LAI, but the audit showed that it duplicated normalized VARI rather than representing an independent estimate of leaf area; it was retained only as a provenance field.

2.2.3. Agronomic Outcomes

At R8, we evaluated grain yield (kg ha^{-1}), plant height, and pod number (Table 1). These outcomes were fitted with the same spatial framework used for the phenomic traits, but they were not included in the primary PCA. Instead, they served as external descriptors of agronomic concordance within ELO_25A (Figure 2).

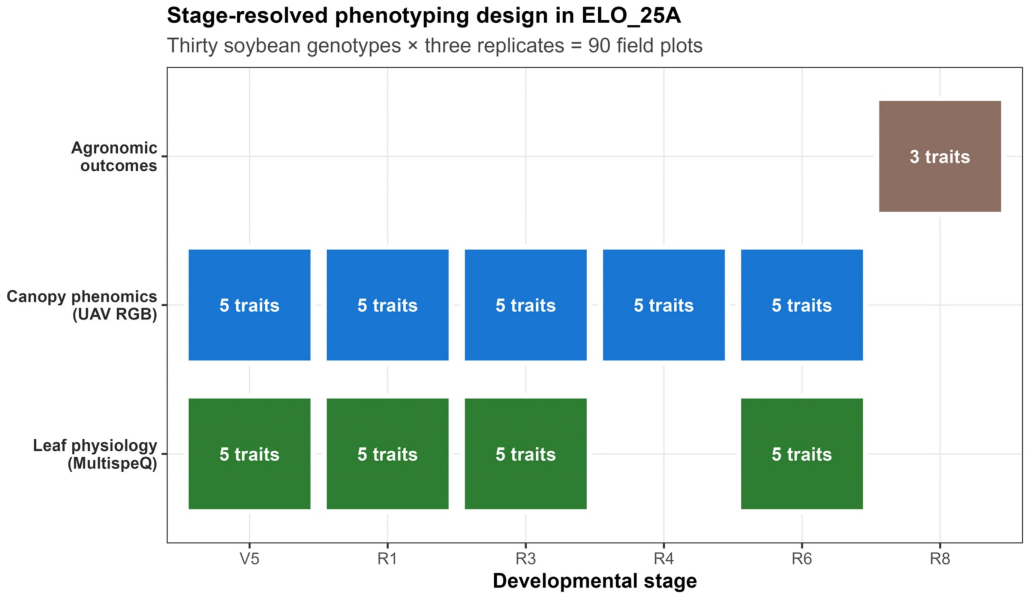


Figure 2. Stage-resolved phenotyping design for 30 soybean genotypes evaluated in ELO_25A. Tiles indicate the number of traits measured by platform at each developmental stage.

Table 1. Experimental phenotyping domains and developmental coverage in ELO_25A.

Domain	Platform	Traits	Stages	Trait–Stage Combinations	Observed Coverage
Leaf physiology	MultispeQ V2.0	ΦPSII, qL, NPQt, Fv′/Fm′, SPAD	V5, R1, R3, R6	20	359/360 plot–stages
Canopy phenomics	DJI Air 2 RGB	NGRDI, VARI, GLI, ExG, TGI	V5, R1, R3, R4, R6	25	450/450 plot–stages
Agronomic outcomes	Field measurements	Yield, plant height, pod number	R8	3	90/90 plots

Note. Thirty genotypes were evaluated in three replicates (90 plots). The 48 combinations comprise 45 phenomic and three agronomic combinations.

2.3. Stagewise Spatial Mixed Models

We analyzed each trait–stage combination separately by residual maximum likelihood in ASReml-R 4.2 [28]. Replicate was fitted as a fixed effect, while genotype and incomplete block were random terms. For physiological traits, standardized PAR was also included in the fixed part of the model. We prespecified four residual structures:

independent errors (IID), AR1 along field rows, AR1 along field columns, and separable AR1(row) × AR1(column). A candidate model was considered only if it converged and passed the parameter and covariance diagnostics. Among the admissible candidates, we selected the simplest model within two AIC units of the minimum.

For combination c, the general model was:

$$y_c = X_c\beta_c + Z_{\{bc\}b_c} + Z_{\{gc\}g_c} + \varepsilon_c \quad (1)$$

In this model, β_c contains the fixed effects, b_c represents incomplete block, g_c represents genotype, and ε_c follows the selected residual structure. We first treated genotype as random to estimate variance components and BLUP precision. Under the same selected spatial model, genotype was then fitted as fixed to obtain adjusted genotype means and the full 30 × 30 prediction covariance matrix required for the two-stage analyses.

Cullis generalized heritability was calculated as:

$$H^2_{\text{Culli's}} = 1 - \frac{\text{mean}\left(SED^2\right)}{2\sigma^2_g} \quad (2)$$

Here, mean (SED^2) denotes the mean squared standard error of pairwise genotype differences, and σ^2_g is the genotypic variance [29,30]. We reported accuracy as $H^2_{\text{Culli's}}$. A 50:50 mixture likelihood-ratio test for the genotype-variance boundary was retained as supplementary evidence [31]. Neither its *p*-value nor a single heritability cutoff determined feature inclusion; precision, temporal structure, biological interpretation, and redundancy were considered together.

2.4. Targeted Temporal Genetic Covariance Models

We investigated temporal covariance only when the stagewise audit identified a repeated trait that was both estimable and biologically interpretable. Adjusted genotype means entered two-stage models with stage fitted as fixed and genotype-by-stage fitted as random. Two prespecified weighting schemes carried stagewise prediction uncertainty into these models: the diagonal of the inverse prediction covariance as the primary scheme, and inverse squared adjusted-mean standard errors as the sensitivity scheme.

For SPAD, we first compared pairwise DIAG and homogeneous-correlation (CORH) structures. The V5–R3 CORH estimate reached the upper boundary, so we conducted a targeted comparison between DIAG and an explicit rank-one model (RR1; $G = \Lambda\Lambda'$). RR1 was selected only when it was admissible and improved both AIC and BIC by more than two units under both weighting schemes. For GLI at R3, R4, and R6, DIAG was compared directly with RR1. We selected the admissible structure with the lowest AIC unless DIAG was within two AIC units. Under RR1, |r_g| = 1 follows structurally from the rank-one constraint and was not interpreted as a freely estimated perfect genetic correlation.

2.5. Construction of the Primary Phenomic Panel

Feature selection combined stagewise precision, temporal redundancy, biological interpretation, and correlations among adjusted genotype means. The resulting panel contained six features: NPQt R3, qL R3, ΦPSII R6, SPAD R3, SPAD R6, and GLI R4. SPAD R3 represented the shared V5–R3 dimension because it had greater stagewise precision, whereas SPAD R6 was retained for its distinct late-season information. GLI R4 represented the R3–R4–R6 dimension because it had the highest stagewise heritability within that module; GLI R3 remained a candidate earlier proxy. NPQt R3 was preferred to Fv′/Fm′ R3 because their adjusted genotype means were almost collinear ($r = -0.993$), and NPQt directly describes regulated energy dissipation. qL R3 and NPQt R3 were also correlated ($|r| = 0.783$), so their partial redundancy was examined by repeating the PCA without

qL. The full correlation structure used to assess redundancy is shown in Supplementary Figure S1. ΦPSII R6 retained a late operating-photochemistry measurement alongside the chlorophyll and canopy features. These decisions produced a biologically interpretable panel for exploratory integration, but the selection was partly data-driven and was not independently validated. The subsequent PCA therefore describes the organization of this selected panel within ELO_25A.

2.6. Multivariate Analysis, Uncertainty and Agronomic Concordance

The six adjusted-mean features were standardized and analyzed by equal-weight PCA. Equal weighting was chosen because precision had already influenced the selection of representative features; weighting again by heritability would give that information a second role. Horn parallel analysis used 2000 simulated matrices. A component was retained only when its eigenvalue exceeded the 95th percentile of the corresponding null distribution [32,33]. Parallel-analysis and loading-uncertainty diagnostics are provided in Supplementary Figures S2 and S3.

Uncertainty was propagated with 2000 parametric draws from each feature-specific 30 × 30 prediction covariance matrix. Cross-feature prediction covariances were unavailable, so the propagation retained the full covariance among adjusted genotype means within each feature but assumed independent estimation errors between features. This assumption concerns the errors of the adjusted means, not the relationships among their estimated values, which remained represented in the PCA. The resulting intervals are therefore conditional on the fitted stagewise covariance estimates and the assumption of zero cross-feature error covariance. For each draw, component order and signs were aligned to the observed PCA before intervals were calculated for explained variance, loadings, and genotype scores. The six-feature panel was held fixed during propagation; feature selection was not repeated in each draw. The intervals therefore describe uncertainty conditional on the selected features and do not include uncertainty associated with choosing the panel. Two planned alternatives tested the stability of the result: an H^2-weighted PCA and an equal-weight PCA without qL R3. Both were compared with the primary solution using Procrustes similarity, aligned score correlations, and correlations between genotype-distance matrices. The corresponding PC1 sensitivity comparisons are shown in Supplementary Figure S4.

Pearson and Spearman correlations related PC1-PC3 scores to yield, plant height, and pod number. Two-sided *p*-values were obtained from 5000 permutations, and 95% intervals from 5000 genotype bootstraps. The Benjamini–Hochberg procedure controlled the false-discovery rate separately within each correlation method [34]. These tests describe agreement within ELO_25A; they are not a validation of prediction or indirect selection.

The first three PCs, which accounted for 79.1% of the observed variance, were also examined by k-means for k = 2–5 [35]. We required a candidate grouping to satisfy all prespecified criteria: silhouette width, bootstrap adjusted Rand index, within-cluster consensus, between-cluster consensus, and the proportion of ambiguous assignments. Solutions failing any criterion were kept only as diagnostics.

2.7. Reproducibility and Portable Genetic Matrices

The workflow kept data auditing, stagewise models, temporal models, matrix export, and multivariate uncertainty in separate, traceable steps. Quality-control checks were passed by the 48 selected stagewise models, 1440 genotype BLUPs, 1440 adjusted means, 43,200 adjusted-mean covariance cells, 192 stagewise matrix collections, and eight final temporal matrices. Portable CSV and matrix files were accompanied by manifests and MD5 checksums, so the downstream analyses can be rerun without an active ASReml

license. This separation also makes the fitted genetic information reusable in other analytical frameworks [15,36].

3. Results

3.1. Stage-Dependent Genetic Signal

All 192 prespecified spatial candidates and all 48 trait–stage combinations evaluated under four residual structures converged. An admissible model was therefore selected for every combination. The preferred residual structure varied across traits and stages: IID, row-AR1, column-AR1, and separable AR1 structures were each supported by field patterns. Supplementary Tables S1 and S2 report the complete parameter estimates and candidate-model comparisons.

Genetic precision changed markedly over development (Figure 3). Several physiological traits carried little or no usable signal at R1 yet became estimable at R3 or R6. SPAD rose from 0.198 at V5 and approximately zero at R1 to 0.428 at R3 and 0.465 at R6. At R3, NPQt and qL reached $H^2_{\text{Culli's}}$ values of 0.306 and 0.344, respectively, while ΦPSII reached 0.360 at R6.

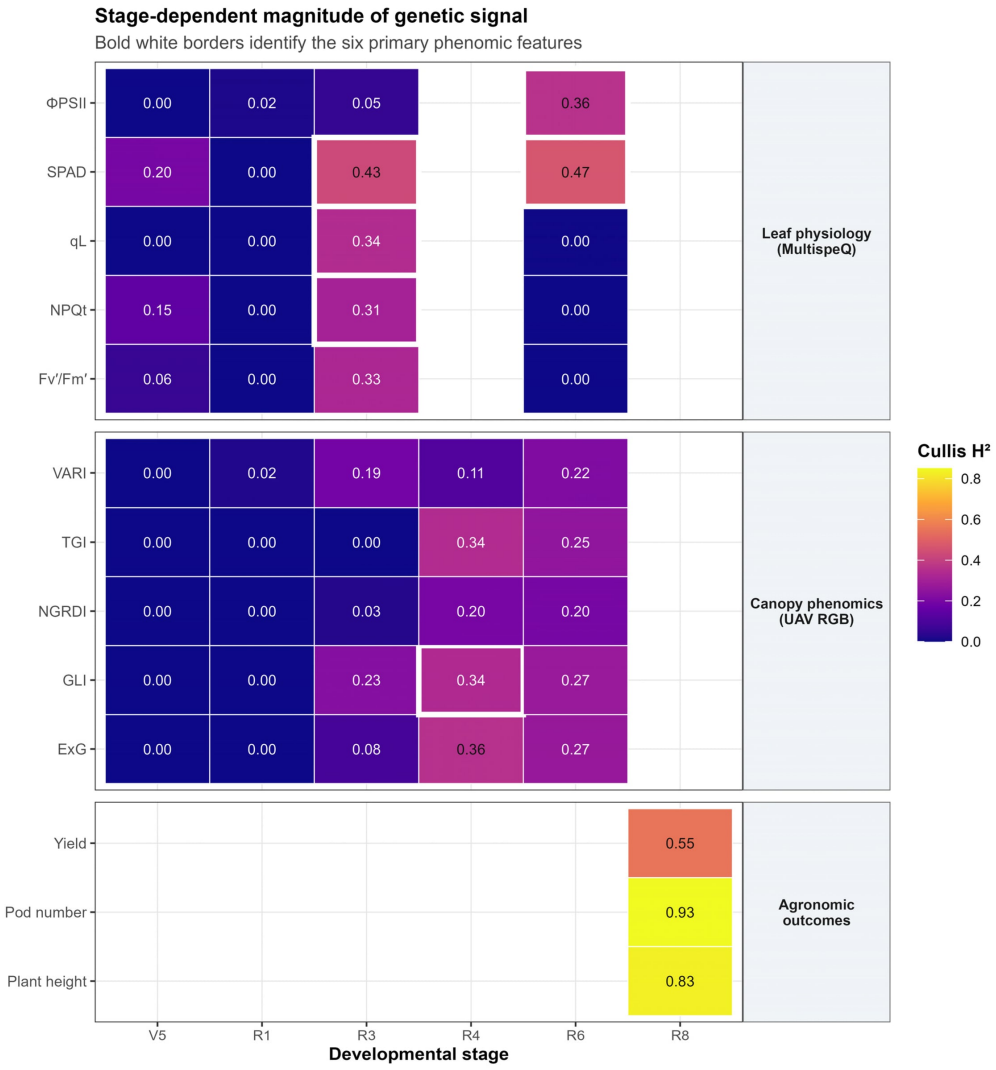


Figure 3. Cullis heritability across 48 trait–stage combinations. Bold white cell borders identify the six features retained for primary multivariate integration.

R6. The UAV traits likewise showed stronger signal during later reproductive stages than at V5 or R1; GLI reached 0.335 at R4. Agronomic outcomes were estimated more precisely, especially pod number ($H^2 = 0.935$) and plant height ($H^2 = 0.831$), whereas yield had moderate heritability ($H^2 = 0.550$).

The six primary phenomic features occupied a narrow but useful range of precision ($H^2 = 0.306$–0.465; accuracy = 0.553–0.682; Table 2). Their retention should not be interpreted as evidence that the other combinations lacked biological relevance. Rather, this set provided a parsimonious balance of estimable signal, complementary biological roles, and nonredundant temporal coverage for this experiment.

Table 2. Primary phenomic features retained for multivariate genetic organization.

Feature	Domain	Biological Dimension	Cullis H^2	Accuracy	Temporal Role
SPAD R6	Leaf physiology	Leaf chlorophyll status	0.465	0.682	Complementary late SPAD state
SPAD R3	Leaf physiology	Leaf chlorophyll status	0.428	0.654	Representative of shared V5–R3 axis
ΦPSII R6	Leaf physiology	Operating photochemistry	0.360	0.600	Stage-specific late dimension
qL R3	Leaf physiology	Photochemical openness	0.344	0.587	Stage-specific R3 dimension
GLI R4	UAV canopy	Canopy greenness	0.335	0.579	Representative of shared R3–R4–R6 axis
NPQt R3	Leaf physiology	Energy dissipation	0.306	0.553	Stage-specific R3 dimension

Note. Heritability and accuracy are stage-specific estimates. They describe precision and were considered with temporal structure and redundancy rather than used as an automatic selection threshold.

3.2. Shared and Stage-Specific Temporal Dimensions

SPAD and GLI showed most clearly that strong stagewise precision can reflect either a shared temporal dimension or stage-specific information (Figure 4; Table 3). For SPAD V5–R3, RR1 was admissible and fit better than DIAG under both weighting definitions (primary $\Delta AIC = \Delta BIC = -7.79$). Under the primary weights, the R3 loading exceeded the V5 loading (2.21 versus 1.80), in line with the stronger expression and greater stagewise precision at R3. Within ELO_25A, these comparisons favored a parsimonious proportional representation of SPAD V5–R3, supporting R3 as the preferred representative and V5 as a candidate earlier proxy. Perfect proportionality follows from the RR1 constraint; it was not independently estimated or validated.

R6, however, was not part of the same supported SPAD dimension. Pairwise models that included SPAD R6 retained DIAG, and adjusted values at R3 and R6 were almost unrelated ($r = 0.019$). We therefore retained R6 as a complementary late chlorophyll state, not as another measurement of the V5–R3 axis.

GLI followed a different pattern. For GLI R3–R4–R6, RR1 was strongly favored over DIAG under both weighting definitions (primary $\Delta AIC = -53.89$; sensitivity $\Delta AIC = -45.92$). Under the selected RR1 structure, similar positive loadings at R3, R4, and R6 summarized the late-canopy measurements through one proportional dimension. Because R4 had the highest stagewise H^2, it was chosen as the primary representative; R3 remains a defensible earlier proxy. This rank-one result denotes shared proportional expression and should not be interpreted as an independently estimated perfect genetic correlation. The selected temporal models and stage loadings are reported in Supplementary Tables S3a–S3c.

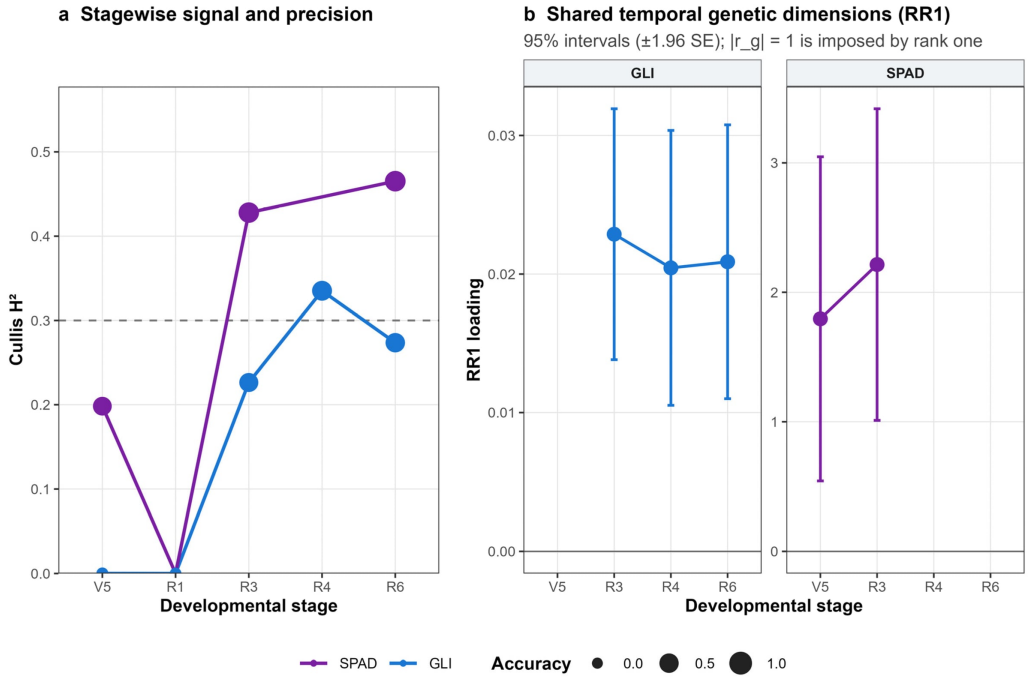


Figure 4. Temporal genetic organization of SPAD and GLI. Stage heritability, shown with prediction accuracy, and RR1 loadings with 95% intervals describe the support for shared genetic dimensions. Under a rank-one model, $|r_g| = 1$ is structurally imposed rather than freely estimated. The dashed horizontal line in panel (**a**) is a visual reference at Cullis $H^2 = 0.30$ and does not represent an automatic selection threshold.

Table 3. Selected temporal genetic models and candidate measurement windows within ELO_25A.

Module	Stages	Selected Structure	ΔAIC vs. DIAG	Primary Window	Interpretation
SPAD early–mid	V5, R3	RR1	−7.79	R3; V5 early proxy	RR1-supported summary with imposed proportionality; greater stagewise precision at R3
SPAD late	R6	Stage-specific	—	R6	Complementary late chlorophyll state, distinct from V5–R3
GLI late canopy	R3, R4, R6	RR1	−53.89	R4; R3 early proxy	RR1-supported late-canopy summary with imposed proportionality

Note. Negative ΔAIC favors RR1 over DIAG. Under RR1, $|r_g| = 1$ is imposed by rank one and is not a freely estimated perfect genetic correlation.

3.3. Multivariate Organization and Propagated Uncertainty

In the equal-weight PCA of the six features, PC1 emerged as the dominant phenomic axis but did not capture all variation. PC1 explained 39.12% of the observed standardized variance and was the only component above the 95th-percentile threshold from Horn parallel analysis. PC2 and PC3 explained 20.70% and 19.31%, respectively, and were retained only for descriptive diagnostics. Together, PC1–PC3 accounted for 79.13%. The complete PCA axis summary and the descriptive PC1-PC2 plane are provided in Supplementary Table S4 and Figure S5, respectively.

PC1 mainly contrasts energy regulation at R3 with leaf chlorophyll status (Figure 5). NPQt R3 and qL R3 had strong positive variable-score correlations (0.922 and 0.803), whereas SPAD R3 was negative (−0.652). The propagated 95% intervals excluded zero for

each of these three loadings. GLI R4, ΦPSII R6, and SPAD R6 added information, but with greater uncertainty, consistent with complementary rather than dominant contributions.

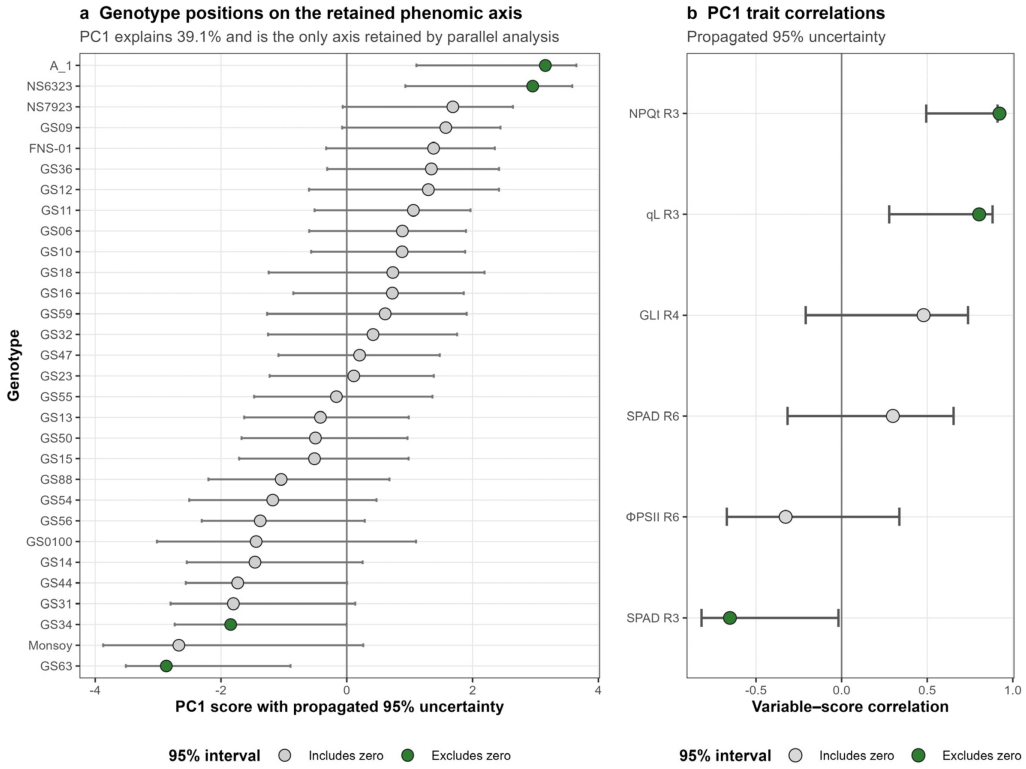


Figure 5. Continuous genotype organization on the only principal component retained by Horn parallel analysis. Score and loading intervals propagate stagewise prediction covariance; no stable categorical cluster solution was supported. The error bars represent 95% confidence intervals calculated as the estimate ± 1.96 SE.

After propagating prediction covariance, PC1 remained identifiable, although separation among individual genotypes was limited. Only A_1 and NS6323 had PC1 intervals entirely above zero; only GS34 and GS63 had intervals entirely below zero. All other genotypes lay along a continuous gradient with overlapping uncertainty. The result supports targeted contrasts among germplasm, but not a complete ranking of all 30 entries.

The PC1 configuration changed little under either prespecified alternative. Relative to the primary analysis, H^2 weighting gave a Procrustes similarity of 0.987 and an absolute PC1-score correlation of 0.994. When qL R3 was removed, the corresponding values were 0.953 and 0.924. Moderate redundancy between qL and NPQt therefore altered the emphasis of the solution but did not generate its dominant configuration.

3.4. Agronomic Concordance and Absence of Stable Clusters

Within ELO_25A, PC1 was associated with plant height and pod number, but not with yield (Figure 6; Table 4). Pearson correlations were 0.550 for plant height and 0.560 for pod number; both remained significant after false-discovery adjustment ($q = 0.006$). The rank-based association was also retained for pod number (Spearman $\rho = 0.499$; $q = 0.041$), whereas the corresponding association with plant height was not ($\rho = 0.421$; $q = 0.082$). Associations with yield were close to zero (Pearson $r = 0.056$; Spearman $\rho = 0.037$; $q \geq 0.822$). Thus, in this experiment, PC1 tracked vegetative architecture and reproductive load without serving as a yield surrogate.

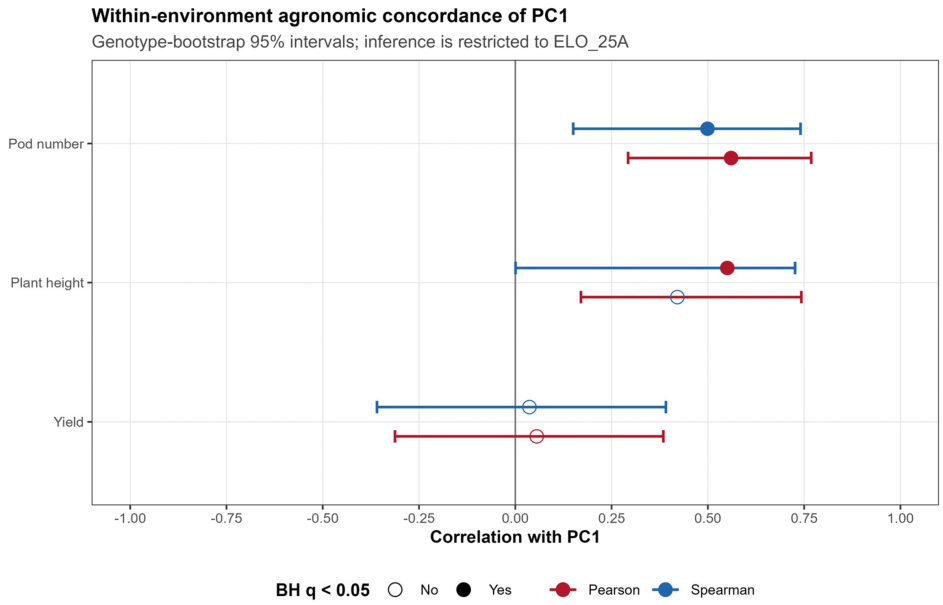


Figure 6. Within-environment concordance between retained PC1 scores and agronomic outcomes. These associations do not constitute predictive validation or indirect-selection validation.

Table 4. Retained multivariate architecture, sensitivity, and agronomic concordance.

Evidence Component	Estimate	Uncertainty or Decision Rule	Interpretation
PC1 variance	39.12%	Propagation interval: 25.10–39.39%	Only axis retained by parallel analysis
Dominant PC1 correlations	NPQt R3: 0.922; qL R3: 0.803; SPAD R3: −0.652	Propagated intervals excluded zero	R3 energy regulation contrasted with chlorophyll status
H^2-weighted sensitivity	Procrustes = 0.987; PC1 |r| = 0.994	Prespecified	Configuration robust to precision weighting
Without qL R3	Procrustes = 0.953; PC1 |r| = 0.924	Prespecified	Configuration robust to redundancy review
Yield concordance	Pearson r = 0.056	95% CI −0.312–0.384; q = 0.822	No within-environment association
Plant-height concordance	Pearson r = 0.550	95% CI 0.170–0.743; q = 0.006	Positive within-environment concordance
Pod-number concordance	Pearson r = 0.560	95% CI 0.293–0.768; q = 0.006	Positive within-environment concordance
Cluster decision	No stable k = 2–5 solution	All stability gates required	Report continuous organization

Note. Agronomic associations are descriptive within ELO_25A. q-values are Benjamini–Hochberg adjusted within the correlation method.

None of the k = 2–5 solutions satisfied every prespecified stability gate. The genotype panel is therefore better represented as a continuous phenomic distribution than as biological classes, subpopulations, or stable clusters. Candidate assignments and consensus matrices are included only as reproducibility diagnostics in Supplementary Table S5 and Figure S6. The complete reproducibility/QC registry and portable file inventory are provided in Supplementary Tables S6 and S7.

4. Discussion

4.1. Stage Dependence Is a Measurement Design Problem

Genetic signal in this study was stage-dependent rather than a fixed property of a universal soybean phenomic panel. Several traits measured with the same instrument showed a negligible signal at R1 and a moderate signal at R3 or R6. This pattern has direct

consequences for experimental design. Restricting phenotyping to an early stage could lead to a breeding program that dismisses useful traits, while measuring every trait at every stage would consume labor and increase multiplicity without a proportional gain in information.

Within ELO_25A, R3–R4 concentrated complementary information: NPQt and qL at R3, the most precise expression of the shared V5–R3 SPAD dimension at R3, and the most precise representation of the shared GLI dimension at R4. This interval therefore provides a candidate core measurement window for the evaluated panel under the conditions of this experiment. Its applicability across maturity groups, tropical environments, and drought conditions remains untested. Referring measurements to reproductive stages provides a developmental reference, but does not ensure that different genotypes reach those stages at the same time or experience comparable conditions. Under drought, the balance between photochemical activity, energy dissipation, chlorophyll status, and canopy development may also differ from the pattern observed here. These considerations could change which traits and stages provide the most useful genotype differentiation. Measurements at V5 and R6 also retained value for earlier assessment and complementary late-season information, respectively. The proposed window should consequently guide validation in connected trials rather than be adopted as a general soybean phenotyping recommendation.

Temporal modeling provided a basis for assessing repeated measurements beyond their stagewise heritability. For SPAD V5–R3 and GLI R3–R4–R6, model comparisons favored RR1 over DIAG under both weighting schemes. These comparisons support a parsimonious summary of the temporal patterns within ELO_25A, although the perfect proportionality of each represented dimension is imposed by rank one. The fitted structures therefore suggest opportunities to reduce repeated measurements rather than establish that the stages are interchangeable. R3 was the more precise SPAD representative, with V5 offering a candidate earlier measurement; R4 was the more precise GLI representative, with R3 offering an earlier alternative. SPAD R6 was retained separately because the examined pairwise comparisons did not support its inclusion in the same proportional summary. Whether these measurement choices preserve useful genotype differentiation elsewhere requires independent validation.

4.2. Complementarity and Redundancy Across Low Sensors

The primary feature panel combined leaf-level physiology with canopy greenness but did not treat every sensor output as an independent trait. That distinction matters because MultispeQ energy-partition variables are mathematically dependent, LEF-related variables derive from ΦPSII and PAR, and several RGB indices were nearly redundant. Including every output would increase dimensionality without a comparable increase in biological information.

NPQt R3 and Fv′/Fm′ R3 provide a clear example of biological and statistical redundancy through their strong inverse relationship. We retained NPQt because it directly describes regulated dissipation; Fv′/Fm′ remained available as a substitute in the sensitivity analysis. qL R3 was also correlated with NPQt, although not to the point of near identity. Removing qL weakened PC1 but did not change its overall configuration. Biological interpretation and sensitivity analyses helped assess redundancy within the selected panel. However, because feature selection and PCA used the same experiment, the resulting configuration remains conditional on those selection decisions. Its similarity under alternative weighting and removal of qL supports consistency under the examined alternatives but does not establish that the six-feature panel is uniquely optimal or that the same organization will emerge in independent data.

Only GLI R4 was needed to represent RGB canopy phenotyping; a suite of highly correlated greenness indices added little. For breeding programs operating with limited

resources, this result shows what may be obtained from a consumer UAV and an open RGB workflow: a stable late-canopy dimension when the measurement stage and feature-reduction rules are chosen carefully. The value came from when and what was measured, rather than from the sheer number of calculated indices.

4.3. Multivariate Architecture Is Continuous and Uncertain

Parallel analysis retained only PC1. Treating PC1–PC3 as equally supported biological axes would therefore exceed the evidence. PC2 and PC3 remain useful descriptive and sensitivity diagnostics, but inference should focus on this axis, which explained 39.1% of the standardized variation. The dominant pattern contrasted high NPQt and qL with high SPAD at R3, consistent with a difference between light-energy regulation and leaf chlorophyll status. Contributions from late ΦPSII, SPAD, and canopy greenness were less certain.

Propagating the adjusted-mean covariance changed how the genotype scores could be interpreted. Point estimates placed all genotypes in order, yet only four entries had intervals that excluded zero. The data therefore support a continuous map of germplasm with explicit uncertainty, not an unqualified ranking. The absence of any cluster solution that met the stability gates leads to the same conclusion. Although k-means will partition any sample, those partitions should not be called biological groups without resampling support.

4.4. Breeding Relevance Without a Yield Prediction Claim

Within ELO_25A, PC1 was associated with plant height and pod number but showed no detectable association with grain yield. This result limits the use of the retained phenomic axis as a yield surrogate, while leaving its value for physiological characterization open. The analysis concerns the composite PC1 score and does not establish that every retained feature is individually unrelated to yield. Its associations with height and pod number connect the phenomic pattern with aspects of plant architecture and reproductive load, but neither association demonstrates a favorable selection response or an improvement in productivity.

The framework can nevertheless inform breeding research by characterizing physiological contrasts among entries and identifying candidate measurements for a more focused phenotyping schedule. These uses concern what to measure and which germplasm contrasts warrant further investigation; they do not identify superior parents or justify selection for higher yield. Before the features or their composite scores guide selection, independent trials should establish whether the observed contrasts and candidate temporal proxies recur across populations and environments. Their contribution to selection decisions should then be evaluated against an agronomic baseline, using validation data separate from feature selection and model development. Claims of indirect selection for yield would additionally require evidence of a relevant genetic association and a resulting selection benefit, neither of which was established here.

4.5. Limitations and Next Validation Steps

Several limitations define the scope of these findings. The experiment included 30 genotypic entries, each represented by three field plots, but replication within the trial does not increase the number of distinct genotypes available for multivariate analysis. The spatial mixed models accounted for the field design, while stagewise heritability estimates described the precision of genotype differentiation under those conditions. That precision varied substantially among traits and stages, so the panel size should not be interpreted as ensuring reliable estimation for every measurement. Reducing the primary PCA to six features limited dimensionality, but the analysis remained exploratory and conditional on the evaluated entries and selected features. Propagated intervals and sensitivity analyses

helped qualify the observed configuration; they did not establish its representativeness beyond this panel. The single location–season combination also limits interpretation of stage dependence. Differences among stages may reflect developmental changes together with the environmental conditions encountered on their respective measurement dates; this experiment cannot isolate those contributions. Consequently, neither the heritability estimates nor the selected temporal structures should be treated as stable properties of these genotypes across tropical environments. Their transferability requires connected trials across locations and seasons that permit estimation of genotype × environment and genotype × environment × stage variation. The two-stage analysis of adjusted genotype means also introduced assumptions that require consideration. An important limitation concerns cross-feature prediction-error covariance, which was unavailable from the separately fitted stagewise models. The propagation preserved covariance among genotype estimates within each feature but treated estimation errors from different features as independent. Although correlations among the adjusted trait values were represented in the PCA, these correlations cannot establish the magnitude or direction of the omitted error covariances. Shared plots, sampling conditions, and measurement procedures could produce correlated errors, particularly among physiological traits measured together. Omitting such dependence may affect the propagated intervals for explained variance, loadings, and genotype scores, but the direction and magnitude of that effect were not established. These intervals should therefore be interpreted as conditional measures of uncertainty rather than as accounting for all sources of estimation error. The weighting and feature-removal sensitivity analyses assessed alternative PCA configurations; they did not test robustness to cross-feature error covariance. Stage categories were available for analysis, but genotype-specific thermal time was not. Differences in developmental timing could therefore contribute to the observed physiological and canopy contrasts, and the analysis cannot establish their independence from maturity-related variation. Because all genotypes were phenotyped on common calendar dates assigned to trial-level stages, individual entries may not have been at precisely the same developmental state at measurement. Accordingly, the relevant limitation is potential within-stage phenological mismatch rather than differential environmental exposure arising from genotype-specific sampling dates. Although field observations suggested a relatively narrow range of phenological development among entries, genotype-specific stage dates and thermal-time records were unavailable; therefore, maturity-related variation could not be fully separated from the observed stage-specific contrasts. In addition, consecutive physiological subsamples within a plot do not provide true within-day temporal replication. The 10:30–14:00 h sampling period restricted the observations to part of the day, but adjustment for recorded PAR did not account for all possible effects of measurement time, recent light exposure, or other environmental conditions. Any systematic correspondence between genotype measurement order and changing conditions could therefore affect genotype comparisons. The physiological results should be interpreted as responses observed within the sampled daytime period, rather than as daily integrated performance.

We limited reduced-rank temporal models to SPAD and GLI after the identifiability audit; broader unstructured longitudinal models were not fitted to the remaining traits. With only 30 genotypes, this constraint avoided unsupported model complexity. Because RR1 imposes perfect proportionality within the represented dimension, the fitted models are parsimonious summaries rather than evidence of an invariant genetic correlation.

Future validation should evaluate the candidate R3–R4 window alongside V5 and R6 in connected trials across locations and seasons, with documented maturity groups and contrasting water availability. Recording genotype-specific stage dates and water conditions would help determine whether the preferred measurement stages persist or shift under different developmental and environmental circumstances. This design would allow esti-

mation of genotype × environment × stage structure, evaluation of the proposed temporal proxies across environments, and testing of whether phenomic information improves selection accuracy over agronomic baselines. Future analyses should also estimate cross-feature prediction-error covariance through a joint modeling framework or a design-preserving resampling procedure that refits the analytical workflow, allowing its consequences for multivariate uncertainty to be evaluated. Where feasible, resampling should repeat spatial-model selection, covariance estimation, and feature selection rather than hold those decisions fixed. This would assess additional uncertainty omitted from the present propagation, while independent validation would remain necessary to evaluate transferability. Operational yield prediction should not be proposed until the findings have been replicated in independent populations. Together, these limitations support treating the framework as an exploratory characterization of ELO_25A with explicitly conditional uncertainty. The reported contrasts and temporal summaries provide hypotheses for further evaluation, but their reproducibility across populations and environments remains unknown. Reliability for breeding decisions will depend on whether those patterns persist under independent sampling and contribute useful information beyond conventional agronomic measurements.

5. Conclusions

Stagewise heritability and temporal covariance captured different aspects of phenomic utility in tropical soybean. The six retained physiological and UAV–RGB features provided complementary genetic information with moderate precision. Temporal model comparisons favored RR1 summaries for SPAD V5–R3 and GLI R3–R4–R6, supporting parsimonious representations within ELO_25A while imposing perfect proportionality through the rank-one constraint. For the evaluated panel in ELO_25A, R3 and R4 were the preferred primary measurement stages based on the observed precision and temporal redundancy. SPAD at V5 and GLI at R3 were candidate earlier proxies within this experiment, whereas SPAD at R6 contributed distinct late-season information. These measurement choices require validation across environments, seasons, and maturity groups before broader use in soybean breeding.

A single phenomic axis, interpreted with its uncertainty, summarized the most reproducible multivariate contrast. It was associated with plant height and pod number, but not with yield, and the genotypes formed a continuum rather than stable clusters. Within ELO_25A, this low-cost phenotyping framework therefore supports germplasm discrimination, efficient measurement scheduling, and generation of testable hypotheses. It does not establish a yield predictor or an indirect-selection criterion.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy16222281/s1, Figures S1–S6: primary-feature correlations, parallel-analysis uncertainty, loading uncertainty, sensitivity of PC1 scores, the descriptive PC1-PC2 plane, and cluster-stability diagnostics. Tables S1 and S2: stagewise genetic parameters and candidate spatial-model comparisons. Tables S3a–S3c: selected temporal models and RR1 loadings for GLI and SPAD. Table S4: PCA axis summary. Table S5: candidate cluster assignments. Tables S6 and S7: reproducibility/QC registry and portable file inventory with checksums.

Author Contributions: J.I.R.-V.: Conceptualization, methodology, management of field trials, data curation, statistical analyses, visualization, and writing—original draft. G.A.M.-R.: Conceptualization, visualization, writing—original draft, and validation. D.F.C.R.: Writing—original draft, conceptualization, resources, and methodology. J.A.B.-B.: Formal analysis, validation, and writing—review and editing. E.E.-M.: Formal analysis, methodology, and validation. M.A.G.D.: Supervision, methodology, and validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Agroinsumos S.A.S.; no grant number was assigned. Academic training and research infrastructure were provided by Universidad Nacional de Colombia.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. For further inquiries, please contact the corresponding author.

Acknowledgments: We thank Salvador Gezan for providing access to ASReml-R (v4.2.0.481), and Agroinsumos S.A.S. for logistical and technical support during soybean breeding and phenotyping activities. Special thanks to the farmers Luis Alberto Lopreto Durán (*Finca La Eloisa*, Obando, Valle del Cauca, Colombia) and Roberto Tisnés Mejía for their valuable collaboration in field management. The authors also thank Héctor Estrada Marín, Diego Alejandro Mancipe Cardenas, Evelyn Ramos Rosero, and Julian Fernando Velazco Amaya for their assistance with data collection.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Furbank, R.T.; Tester, M. Phenomics—Technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [CrossRef] [PubMed]
2. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [CrossRef] [PubMed]
3. Reynolds, M.; Chapman, S.; Crespo-Herrera, L.; Molero, G.; Mondal, S.; Pequeno, D.N.; Pinto, F.; Pinera-Chavez, F.J.; Poland, J.; Rivera-Amado, C.; et al. Breeder friendly phenotyping. *Plant Sci.* **2020**, *295*, 110396. [CrossRef] [PubMed]
4. Baker, N.R. Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo. *Annu. Rev. Plant Biol.* **2008**, *59*, 89–113. [CrossRef] [PubMed]
5. Kuhlgert, S.; Austic, G.; Zegarac, R.; Osei-Bonsu, I.; Hoh, D.; Chilvers, M.I.; Roth, M.G.; Bi, K.; TerAvest, D.; Weebadde, P.; et al. MultispeQ Beta: A Tool for Large-Scale Plant Phenotyping Connected to the Open PhotosynQ Network. *R. Soc. Open Sci.* **2016**, *3*, 160592. [CrossRef]
6. Tietz, S.; Hall, C.C.; Cruz, J.A.; Kramer, D.M. $NPQ_{(T)}$: A chlorophyll fluorescence parameter for rapid estimation and imaging of non-photochemical quenching of excitons in photosystem-II-associated antenna complexes. *Plant Cell Environ.* **2017**, *40*, 1243–1255. [CrossRef] [PubMed]
7. Borra-Serrano, I.; De Swaef, T.; Quataert, P.; Aper, J.; Saleem, A.; Saeys, W.; Somers, B.; Roldán-Ruiz, I.; Lootens, P. Closing the Phenotyping Gap: High Resolution UAV Time Series for Soybean Growth Analysis Provides Objective Data from Field Trials. *Remote Sens.* **2020**, *12*, 1644. [CrossRef]
8. Hu, J.; Yue, J.; Xu, X.; Han, S.; Sun, T.; Liu, Y.; Feng, H.; Qiao, H. UAV-Based Remote Sensing for Soybean FVC, LCC, and Maturity Monitoring. *Agriculture* **2023**, *13*, 692. [CrossRef]
9. Moreira, F.F.; Hearst, A.A.; Cherkauer, K.A.; Rainey, K.M. Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods* **2019**, *15*, 139. [CrossRef] [PubMed]
10. Roth, L.; Barendregt, C.; Bétrix, C.-A.; Hund, A.; Walter, A. High-throughput field phenotyping of soybean: Spotting an ideotype. *Remote Sens. Environ.* **2022**, *269*, 112797. [CrossRef]
11. Escamilla, D.M.; Huang, M.; McHale, L.; Wang, D.; Diers, B.; Xavier, A.; Rainey, K.M. Canopy coverage phenotyping and field spatial variability adjustment as an efficient selection tool in soybean breeding. *Crop Sci.* **2023**, *63*, 3277–3291. [CrossRef]
12. Alves, A.K.S.; Araújo, M.S.; Chaves, S.F.S.; Dias, L.A.S.; Corrêdo, L.P.; Pessoa, G.G.F.A.; Bezerra, A.R.G. High throughput phenotyping in soybean breeding using RGB image vegetation indices based on drone. *Sci. Rep.* **2024**, *14*, 32055. [CrossRef] [PubMed]
13. Pawar, P.S.; Wu, C.; Harrison, D.; Rogers, D.; Fallen, B.; Canella Vieira, C. UAV-based phenotyping outperforms visual canopy wilting for evaluating soybean drought tolerance and yield retention under rainfed conditions. *Front. Plant Sci.* **2026**, *17*, 1835549. [CrossRef] [PubMed]
14. Araus, J.L.; Kefauver, S.C.; Zaman-Allah, M.; Olsen, M.S.; Cairns, J.E. Translating High-Throughput Phenotyping into Genetic Gain. *Trends Plant Sci.* **2018**, *23*, 451–466. [CrossRef] [PubMed]
15. Crossa, J.; Fritsche-Neto, R.; Montesinos-Lopez, O.A.; Costa-Neto, G.; Dreisigacker, S.; Montesinos-Lopez, A.; Bentley, A.R. The Modern Plant Breeding Triangle: Optimizing the Use of Genomics, Phenomics, and Enviromics Data. *Front. Plant Sci.* **2021**, *12*, 651480. [CrossRef] [PubMed]
16. Valencia, R.A.; Ligarreto, G.A. *Mejoramiento Varietal y Potencial Genético de la Soya en el Piedemonte Llanero*; Corporación Colombiana de Investigación Agropecuaria—AGROSAVIA: Bogota, Colombia, 2010.
17. Valencia-Ramírez, R.A.; Caicedo, S.; Reyes, J.C.; Flórez-Gómez, D.L.; Tibocha, Y. Genetic Parameters of Agronomic Traits and Yield Components of Soybean Genotypes in the Colombian Altillanura. *Rev. U.D.C.A Act. Divulg. Cient.* **2022**, *25*, e1901. [CrossRef]
18. Valencia Ramírez, R.A.; Tibocha Ardila, Y.S. Application of the BLUPe Predictor in the Selection of Potential Soybean Varieties for Orinoquia. *Agron. Colomb.* **2022**, *40*, 165–173. [CrossRef]

19. Sparks, A. nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *J. Open Source Softw.* **2018**, *3*, 1035. [CrossRef]
20. Fehr, W.R.; Caviness, C.E. *A Visual Guide to Soybean Growth Stages*; Cooperative Extension Service, Agriculture and Home Economics Experiment Station, Iowa State University: Ames, IA, USA, 1977.
21. Qin, S.; Ding, Y.; Zhou, Z.; Zhou, M.; Wang, H.; Xu, F.; Yao, Q.; Lv, X.; Zhang, Z.; Zhang, L. Study on the nitrogen content estimation model of cotton leaves based on "image-spectrum-fluorescence" data fusion. *Front. Plant Sci.* **2023**, *14*, 1117277. [CrossRef] [PubMed]
22. Ali, W.; Grzybowski, M.; Torres-Rodríguez, J.V.; Li, F.; Shrestha, N.; Mathivanan, R.K.; de Bernardeaux, G.; Hoang, K.; Mural, R.V.; Roston, R.L.; et al. Quantitative genetics of photosynthetic trait variation in maize. *J. Exp. Bot.* **2025**, *76*, 4141–4153. [CrossRef] [PubMed]
23. Rife, T.W.; Poland, J.A. Field Book: An Open-Source Application for Field Data Collection on Android. *Crop Sci.* **2014**, *54*, 1624–1627. [CrossRef]
24. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; Mcmurtrey, J.E.; Walthall, C.L. Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. *Precis. Agric.* **2005**, *6*, 359–378. [CrossRef]
25. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [CrossRef]
26. Louhaichi, M.; Borman, M.M.; Johnson, D.E. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. *Geocarto Int.* **2001**, *16*, 65–70. [CrossRef]
27. Woebbecke, D.M.; Meyer, G.E.; Von Bargen, K.; Mortensen, D.A. Color Indices for Weed Identification Under Various Soil, Residue, and Lighting Conditions. *Trans. ASAE* **1995**, *38*, 259–269. [CrossRef]
28. Butler, D.G.; Cullis, B.R.; Gilmour, A.R.; Gogel, B.J.; Thompson, R. *ASReml-R Reference Manual Version 4.2*; VSN International Ltd.: Hemel Hempstead, UK, 2023.
29. Cullis, B.R.; Smith, A.B.; Coombes, N.E. On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* **2006**, *11*, 381–393. [CrossRef]
30. Piepho, H.-P.; Möhring, J. Computing Heritability and Selection Response From Unbalanced Plant Breeding Trials. *Genetics* **2007**, *177*, 1881–1888. [CrossRef] [PubMed]
31. Self, S.G.; Liang, K.-Y. Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions. *J. Am. Stat. Assoc.* **1987**, *82*, 605–610. [CrossRef]
32. Horn, J.L. A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika* **1965**, *30*, 179–185. [CrossRef] [PubMed]
33. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [CrossRef] [PubMed]
34. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1995**, *57*, 289–300. [CrossRef]
35. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* **1979**, *28*, 100. [CrossRef]
36. Crain, J.; Mondal, S.; Rutkoski, J.; Singh, R.P.; Poland, J. Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *Plant Genome* **2018**, *11*, 170043. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.