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.

Environmental and Agronomic Factors Associated with Yield Variation in Two Soybean Cultivars across Multiple Locations

Jung Min Kim
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Jun Hoi Kim
itomi123@korea.kr

Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Beom-Kyu Kang
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Jeong-Hyun Seo
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Su-Vin Heo
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Ga-Eun Kim
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Jihee Park
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Hye Rang Park
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Yo-Han Yoo
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Ye-Jin Son

Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Won Young Han
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Min-Jung Seo
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

Nam Geol Kim
Planning and Coordination Division, National Institute of Crop and Food Science, Rural Development Administration

Seuk Ki Lee
Central-Northern Region Crop Research Center, National Institute of Crop and Food Science, Rural Development Administration

In Hye Lee
Central-Northern Region Crop Research Center, National Institute of Crop and Food Science, Rural Development Administration

Joon-Tae Kim
Gangwon-do Agricultural Research & Extension Service

Jinsil Choi
Jeollanam-do Agricultural Research & Extension Service

Jong-Min Ko
Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration

---

Research Article

Keywords: soybean, yield, Location × Year × Cultivar interaction, multiple regression

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11544176/v1

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

Additional Declarations: No competing interests reported.

Environmental and Agronomic Factors Associated with Yield Variation in Two Soybean Cultivars across Multiple Locations

Running title: Yield Variation in Two Soybean Cultivars across Locations

Jung Min Kim[1], Jun Hoi Kim[1]*, Beom-Kyu Kang[1], Jeong-Hyun Seo[1], Su-Vin Heo[1], Ga-Eun Kim[1], Jihee Park[1], Hye Rang Park[1], Yo-Han Yoo[1], Ye-Jin Son[1], Won Young Han[1], Min-Jung Seo[1], Nam Geol Kim[2], Seuk Ki Lee[3], In Hye Lee[3], Joon-Tae Kim[4], Jinsil Choi[5], Jong-Min Ko[1]

[1] Upland Crop Breeding Division, Department of Upland Crop Sciences, National Institute of Crop and Food Science, Rural Development Administration, Miryang 50424, Republic of Korea

[2] Planning and Coordination Division, National Institute of Crop and Food Science, Rural Development Administration, Wanju 55365, Republic of Korea

[3] Central-Northern Region Crop Research Center, National Institute of Crop and Food Science, Rural Development Administration, Suwon 16429, Republic of Korea

[4] Gangwon-do Agricultural Research & Extension Service, Chuncheon 24203, Republic of Korea

[5] Jeollanam-do Agricultural Research & Extension Service, Naju 58213, Republic of Korea

*Corresponding author: Jun Hoi Kim. E-mail: itomi123@korea.kr, Tel: +82-55-350-1235

Keywords: soybean, yield, Location × Year × Cultivar interaction, multiple regression

Acknowledgements The authors thank the staff of the experimental stations for their assistance with field management and data collection, and the Rural Development Administration for its support of this research.

Declarations

Funding This research was supported by the Rural Development Administration (Project No. RS-2024-00397381).

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval and consent Not applicable

Data availability The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Author contribution Conceptualization: Ko JM, Kim JH; Methodology and formal analysis: Kang BK, Kim JM; Investigation: Seo JH, Heo SV, Park J, Yoo YH, Han WY, Kim NG, Lee SK, Lee IH, Kim JT, Choi J; Data curation: Kim GE, Son YJ; Visualization: Kim JM; Writing–original draft: Kim JM; Writing–review and editing: Park HR, Seo MJ; Supervision and funding acquisition: Kim JH. All authors read and approved the final manuscript.

Abstract

Soybean (*Glycine max* (L.) Merr.) yield reflects interactions among climatic conditions, soil properties, and agronomic traits that vary among cultivars. This study evaluated the medium-late-maturing cultivar 'Seonpung' and the medium-maturing cultivar 'Seonyu2' across eight locations in South Korea over two years. Representative variables were selected by hierarchical clustering based on Pearson correlation and principal component analysis, followed by multiple regression of yield and three yield-related traits: pod number, branch number, and hundred-seed weight. Significant Location × Cultivar ($F = 4.51$, $p < 0.01$) and Location × Year × Cultivar ($F = 7.51$, $p < 0.001$) interactions were detected for yield. Among the 12 yield models, only the Seonpung agronomic model was significant (adjusted $R^2 = 0.51$), in which days after sowing to R6 was negatively associated with yield ($B = -18.97$). In contrast, seven of the 12 pod number models were significant, mainly in Seonpung and the combined dataset (adjusted R^2 values up to 0.45), indicating pod number was more consistently associated with environmental and agronomic factors than final yield. Theoretical yield based on seed number, hundred-seed weight, and plant density was not significantly correlated with observed plot yield ($r = 0.20$; ratio range = 0.54–6.42), suggesting individual-plant measurements may not fully capture plot-scale yield variation. Overall, environmental and agronomic factors were more clearly associated with yield-related traits, particularly pod number, than with final yield, which reflected more complex cultivar-specific responses across locations and years.

Introduction

Recent climate change has increased crop exposure to high temperatures, drought, excessive rainfall, and other extreme weather events that can reduce soybean growth and yield (Gonçalves et al. 2021; Lesk et al. 2022). Meta-analytic estimates indicate that each 1 ℃ increase in global mean temperature would reduce soybean yield by approximately 3.1% (Zhao et al. 2017). In South Korea, soybean is a strategic crop for food security, yet the self-sufficiency rate remains low, although it rose from 28.6% in 2022 to 35.7% in 2023. To expand domestic production, the Strategic Crop Direct Payment Program has promoted soybean cultivation on rice paddies. As a result, the national soybean cultivation area reached approximately 74,000 ha and production approximately 156,000 tons in 2025 (KOSIS 2025; MAFRA 2025). Production is concentrated in a few provinces: Jeollabuk-do (34.5%), Jeollanam-do (12.0%), and Gyeongsangbuk-do (11.6%) together accounting for more than half of national output. As cultivation expands into more diverse environments, understanding the environmental and agronomic factors associated with yield variation is essential for improving yield stability and cultivar selection.

Soybean growth and development are strongly regulated by temperature and photoperiod. As a quantitative short-day crop, soybean responds to these factors through changes in vegetative and reproductive growth (plant height, node number, flowering time, and maturity) (Nakamichi 2015; Wang et al. 2020). Adaptation to latitude is closely related to flowering and maturity characteristics, and cultivars grown outside their adapted areas may show altered developmental duration, morphology, and yield (Wang et al. 2020). Such responses are particularly relevant in South Korea, which has considerable climatic variation despite its small land area. Long-term meteorological records show a north–south temperature gradient, marked spatial variation in precipitation, and warming trends, especially in spring and summer (Jeung et al. 2019; Rahman et al. 2025). The national mean annual temperature for the 1991–2020 normal period was 12.8 ℃, approximately 0.3 ℃ higher than that for 1981–2010, and precipitation decreased in parts of central Korea but increased in southern regions (KMA 2021). Soybean cultivars grown at different locations may therefore experience substantially different thermal and moisture conditions during vegetative and reproductive development. Differences in sowing date can further shift the developmental stage at which plants encounter these conditions, contributing to variation in crop growth and yield.

Soil properties and agronomic management also contribute to differences in soybean production. Soil physicochemical properties of Korean upland fields, including pH, available phosphate, and electrical conductivity, vary among locations (Kong et al. 2015; Roh et al. 2018; Kim et al. 2019; Yoon et al. 2020). These properties can affect nutrient availability, root development, water relations, and thereby crop productivity. Differences in soil chemical properties have been associated with soybean yield and seed quality (Anthony et al. 2012). Agronomic management can also modify yield responses. In particular, plant density and planting configuration influence canopy development, light interception, branching, and the

formation of reproductive structures (De Bruin and Pedersen 2009; Cox et al. 2010; Cox and Cherney 2011). Soybean can compensate for changes in plant density, with adjustments in branch number, pod number, and seed number per plant helping to maintain yield (Cox et al. 2010). Climatic, soil, and agronomic conditions should therefore be considered together when evaluating soybean performance across production environments.

Reproductive development and the formation of yield components are also closely related to soybean yield. The duration of seed filling has long been associated with seed yield across soybean germplasm and breeding populations (Nelson 1986; Smith and Nelson 1986a, 1986b), probably because a longer effective seed-filling period allows more time for dry matter accumulation in developing seeds (Egli 2004). However, final yield is the integrated outcome of multiple components rather than the response of a single trait. Seed number per unit area and individual seed weight are major determinants of soybean yield, and seed number in turn depends on pod number and seeds per pod (Vogel et al. 2021). These components are established at different phenological stages and may differ in their sensitivity to environmental and agronomic conditions. Soybean can also compensate among yield components, which may weaken direct associations between individual environmental factors and final yield. Therefore, evaluating yield components together with final yield may provide additional insight into cultivar-specific responses during yield formation.

'Seonpung' and 'Seonyu2' are Korean soybean cultivars that differ in maturity and production characteristics. Seonpung is a medium-late-maturing, high-yielding cultivar developed for soybean paste and tofu production, and it performed well in regional yield trials (Kim et al. 2017). Seonyu2 is a medium-maturing cultivar developed for double-cropping systems and is suitable for cultivation after winter crops such as wheat, barley, and onion (Kang et al. 2022). Seo et al. (2024) evaluated the responses of both cultivars to sowing date and planting population under southern Korean conditions and reported cultivar-dependent changes in growth and yield. However, comparative information on the performance of these cultivars across diverse soybean-growing locations, and on the environmental and agronomic factors associated with yield variation, remains limited. This study therefore evaluated Seonpung and Seonyu2 at eight locations in South Korea over two consecutive years. The objectives were to (i) characterize variation in agronomic traits, phenological stages, yield and evaluate location × year × cultivar interactions; (ii) identify climatic, soil, and agronomic factors associated with yield variation in each cultivar using correlation-based hierarchical clustering, principal component analysis, and multiple regression; (iii) determine whether these factors showed similar relationships with major yield-related traits; and (iv) examine the correspondence between component-based yield estimates and observed plot yield.

Materials and Methods

Plant materials and field trials

Two soybean cultivars that differ in maturity, the medium-late-maturing 'Seonpung' and the medium-maturing 'Seonyu2', were used in this study (Kim et al. 2017; Kang et al. 2022). Field trials were conducted at eight locations in South Korea (Buan, Dangjin, Goesan, Naju, Paju, Sangju, Suwon, and Youngwol) over two consecutive years, 2024 and 2025. At each location, both cultivars were grown under locally adopted field management practices. Sowing dates varied among locations and years, ranging from May 28 to June 27 in 2024 and from June 5 to June 25 in 2025 (Table S1). Each experimental plot consisted of 4-m-long rows. Planting configurations, including row spacing, within-row spacing, and the number of plants per hill, varied among location × year × cultivar combinations according to local cultivation practices and field conditions. Other crop management practices followed standard practices at each experimental location.

Environmental, soil, and agronomic measurements

Climatic variables included minimum, maximum, and mean air temperature, rainfall, sunshine duration, and solar radiation. Rainfall, sunshine duration, and solar radiation data were additionally compared with records from the Korea Meteorological Administration (KMA). Soil samples collected from each experimental location were analyzed for pH, total nitrogen (TN), organic matter (OM), available phosphate (P_2O_5), and electrical conductivity (EC).

Agronomic traits included plant height (PH), first-pod height (FPH), node number (NN), branch number (BN), pod number (PN), seed number (SN), hundred-seed weight (HSW), plant density (PD), and yield. NN, BN, PN, and SN were measured on a per-plant basis. Seeds per pod (SPP) was calculated as SN divided by PN. Phenological development was recorded as days after sowing (DAS) to the R2, R4, R6, and R8 growth stages, corresponding to full bloom, full pod, full seed, and full maturity, respectively (Fehr and Caviness 1977). These variables were designated as r2d, r4d, r6d, and r8d, respectively. Agronomic and phenological traits were evaluated for each location × year × cultivar combination.

Because planting configurations differed among experimental locations, PD was calculated for each experimental plot using row spacing, within-row spacing, and the number of plants per hill. PD was calculated as follows:

$$\text{Plant density (plants } 10\text{a}^{-1}) = \frac{1000}{\text{row spacing (m)} \times \text{within} - \text{row spacing(m)}} \times \text{plants per hill}$$

For presentation in the tables, plant density was expressed as 10^3 plants 10 a^{-1}.

Statistical analysis

Analysis of variance and multiple comparisons

Trait variability was expressed as the coefficient of variation (CV). The effects of location (L), year (Y), and cultivar (C) on each trait were evaluated using Type III analysis of variance (ANOVA). Two-way ANOVA (L × Y) was conducted separately for each cultivar, whereas three-way ANOVA (L × Y × C) was

applied to the combined dataset. For climatic variables lacking within-plot replication, the L × Y interaction could not be estimated. Mean comparisons between 2024 and 2025 were performed within each location using the Tukey–Kramer HSD test at $\alpha = 0.05$.

Variable selection and regression analysis

To minimize multicollinearity among candidate predictors, variables were grouped by hierarchical clustering based on Pearson correlations (distance = $1 - |r|$; clustering threshold $|r| \geq 0.5$; Cohen 1988). Within each cluster, the variable with the highest loading on the first principal component was selected as the representative variable, and variance inflation factors (VIF) were subsequently calculated to confirm low multicollinearity (VIF < 10). Variables not selected as cluster representatives, such as r6d and sunshine duration, were also retained when they were significantly associated with the dependent variable in the corresponding full models (Holm-adjusted $p < 0.05$) and had acceptable VIF values.

Multiple linear regression was conducted for yield and three yield-related traits, pod number (PN), branch number (BN), and hundred-seed weight (HSW), separately for each cultivar and the combined dataset (Table S2). Models were constructed using four predictor categories: climate, soil, agronomic, and combined climate + soil. To reduce multicollinearity, highly correlated variables were excluded from the predictor pool when related traits were used as dependent variables, and representative predictors were re-selected from the remaining variables. Climate, soil, and combined climate + soil predictor sets were maintained consistently across dependent variables within each dataset.

To evaluate the correspondence between component-based yield estimates and plot-level yield, theoretical yield was calculated as follows:

$$\text{Theoretical yield (kg } 10a^{-1}) = \text{SN} \times \left(\frac{\text{HSW}}{100}\right) \times \frac{\text{plant density}}{1{,}000}$$

where SN is seed number per plant, HSW is hundred-seed weight (g), and plant density is expressed as plants 10 a^{-1}. The theoretical yield was compared with observed plot yield using Pearson correlation. All statistical analyses were performed in R version 4.6.0.

Results

Variation in agronomic traits and environmental conditions

Considerable variation was observed among agronomic traits across locations, years, and cultivars (Tables 1 and S3). Seed number (SN) showed the highest coefficient of variation (CV, 70.7%), followed by pod number (PN, 61.4%), first-pod height (FPH, 55.5%), and branch number (BN, 55.3%). In contrast, phenological traits were relatively less variable, with CVs ranging from 8.3% for r6d to 15.4% for r2d. Yield showed an intermediate level of variation (CV = 37.9%).

Environmental variables also differed in their degree of variability (Tables 2 and S4; Figure S1). Mean

and maximum air temperatures were relatively stable (CV = 2.8% and 5.1%, respectively), whereas minimum temperature and rainfall showed greater variation (31.7% and 26.4%); sunshine duration and solar radiation showed intermediate variation (14.4% and 12.0%). Among soil properties, available phosphate (P_2O_5) and electrical conductivity (EC) exhibited particularly high variability, with CVs of 146.0% and 114.8%, respectively, followed by organic matter (OM; 80.5%) and total nitrogen (TN; 39.7%). Much of this variability reflected differences between years: mean OM was 17.07% in 2024 and 2.24% in 2025, P_2O_5 was 14.65 and 426.34 mg kg^{-1}, EC was 0.14 and 0.64 dS m^{-1}, and TN was 0.08% and 0.14%, respectively (Table 2).

Effect of location, year, and cultivar on agronomic traits and yield

The combined analysis of variance revealed significant effects of location (L), year (Y), and cultivar (C) on soybean yield ($F = 13.11$, $p < 0.001$; $F = 10.87$, $p < 0.01$; and $F = 37.27$, $p < 0.001$, respectively; Table 3). Significant L × Y ($F = 23.17$, $p < 0.001$), L × C ($F = 4.51$, $p < 0.01$), and L × Y × C ($F = 7.51$, $p < 0.001$) interactions were also observed for yield, whereas the Y × C interaction was not significant.

Cultivar-specific analyses showed contrasting patterns of year variation. In Seonpung, significant year effects were observed for plant height (PH), first-pod height (FPH), pod number (PN), seed number (SN), days after sowing to R6 (r6d), hundred-seed weight (HSW), and yield. In Seonyu2, year effects were significant for all agronomic and phenological traits, except HSW. The L × Y interaction for yield was significant in both cultivars, with a larger F-value in Seonyu2 ($F = 61.20$, $p < 0.001$) than in Seonpung ($F = 5.83$, $p < 0.001$), indicating that the difference in yield between years varied more among locations in Seonyu2 (Figure 1).

Environmental and soil variables also exhibited significant spatial and temporal variation (Table S5). Location effects were significant for all measured climatic and soil variables except minimum temperature. Year effects were significant for maximum temperature and for all soil properties, but not for the other climatic variables in the combined analysis. Significant L × Y interactions were detected for pH, EC, OM, and P_2O_5 in both cultivars and the combined dataset, and for TN in Seonyu2 and the combined dataset. Interactions among climatic variables could not be estimated because of the lack of within-plot replication.

Correlation analysis and variable selection

Candidate climatic, soil, and agronomic variables showed several groups of strongly correlated variables (Figure S2). For example, the phenological stages r2d, r4d, r6d, and r8d were strongly correlated with each other (r = 0.77–0.88), as were sunshine duration and solar radiation (r = 0.89) and EC and P_2O_5 (r = 0.87). Among yield-related traits, pod number (PN), seed number (SN), branch number (BN), and seeds per pod (SPP) were moderately to strongly correlated with each other ($|r| = 0.48$–0.98). Because including these traits together as predictors would exceed acceptable multicollinearity levels (e.g., VIF = 36.4 and

43.6 for PN and SN, respectively, when entered jointly in the combined dataset), they were not entered together: when one of them was the dependent variable, the related traits were excluded from the predictor pool (Table S2).

Hierarchical clustering based on Pearson correlation grouped highly correlated variables into clusters (Figure S2), and within each cluster, the variable with the highest loading on the first principal component (PC1) was selected as the representative variable (Figure S3). The selected representative variables showed acceptable multicollinearity levels (VIF < 10) and were subsequently used as predictors in multiple regression models (Table S2).

Multiple regression analysis of yield

Among the 12 yield regression models constructed across three datasets and four predictor categories, only the Seonpung agronomic model was statistically significant (adjusted $R^2 = 0.51$, $F = 4.37$, $p = 0.007$; Tables 4 and S6; Figure 2). In this model, r6d showed the strongest negative association with yield ($B = -18.97$, $p < 0.001$), and r6d and r8d were the only predictors that remained significant after Holm correction (Table S6; Figure 3a). None of the yield models for Seonyu2 or the combined dataset reached overall statistical significance.

The relationship between r6d and yield was further examined by grouping observations according to r6d. In Seonpung, mean yield decreased across increasing r6d terciles, from 382.6 kg 10 a^{-1} in the shortest-r6d group (91–98 days) to 264.5 kg 10 a^{-1} in the longest-r6d group (105–112 days), consistent with the negative association identified in the regression analysis (Figure 3a). In Seonyu2, no comparable linear pattern was observed.

Multiple regression analysis of yield-related traits

As the yield models generally showed limited explanatory ability, the same regression framework was applied to three yield-related traits, PN, BN, and HSW, to determine whether these traits showed clearer associations with the measured climatic, soil, and agronomic factors. Because PN, SN, BN, and SPP were moderately to strongly correlated ($|r| = 0.48–0.98$), the other members of this correlated group were excluded from the candidate predictor pool when PN or BN was used as the dependent variable, and representative predictors were re-selected from the remaining variables using the same clustering and PCA-based procedure described above (Table S2).

Among the three yield-related traits, PN showed the most consistent model-level associations, with 7 of the 12 models reaching overall statistical significance across climate, soil, and agronomic predictor categories in Seonpung and the combined dataset (Figure 2; Tables 4 and S6). In Seonpung, the soil model was significant (adjusted $R^2 = 0.40$, $p = 0.002$), with available phosphate (P_2O_5) positively associated with PN ($B = 0.058$, $p = 0.001$). The Seonpung climate model was also significant (adjusted $R^2 = 0.45$, $p = 0.004$),

with rainfall positively associated with PN (B = 0.069, p = 0.001; Figure 3b); the climate model for the combined dataset showed the same pattern (adjusted R^2 = 0.36, $p < 0.001$; rainfall B = 0.061, $p < 0.001$). In the Seonpung agronomic model (adjusted R^2 = 0.44, p = 0.007), PN was positively associated with node number (B = 11.95, p = 0.005) and negatively associated with plant density (B = -0.005, $p < 0.001$) and plant height (B = -1.03, p = 0.040).

BN showed a similar but less consistent pattern, with 6 of the 12 models reaching overall significance. In contrast, only one of the 12 HSW models was significant, namely the climate + soil model for the combined dataset (adjusted R^2 = 0.24, p = 0.017). None of the regression models for PN, BN, or HSW reached overall significance in Seonyu2.

Correspondence between theoretical and observed yield

Given that individual yield-related traits, particularly pod number, were more consistently associated with environmental and agronomic factors, the weak explanatory power observed for final yield was unexpected under a simple multiplicative yield-component model. To examine this discrepancy directly, a theoretical yield estimate was reconstructed from measured seed number, HSW, and plant density and compared with observed plot yield (Figure S4). The reconstructed estimate was not significantly correlated with observed yield (r = 0.20, p = 0.21, n = 42), and the ratio of theoretical to observed yield ranged from 0.54 to 6.42 across observations. This wide and inconsistent divergence suggests that the weak relationship between yield-related traits and final yield may not simply reflect an absence of biological linkage, but may instead reflect a gap between plant-level sample measurements and plot-level harvested yield, potentially arising from within-plot plant-to-plant variability, harvest losses not captured in hand-sampled measurements, or deviation of observed stand density from nominal plant density.

Discussion

The present study showed that soybean yield and yield-related traits responded differently across locations, years, and cultivars. Significant interactions among location, year, and cultivar for yield indicated that cultivar performance depended on the combination of growing location and year. Yield was lower in 2025 than in 2024 in both cultivars, a year with higher mean maximum temperature (38.7 vs. 35.8 °C; Table 2), which is consistent with global estimates that each 1 °C increase in mean temperature would reduce yields of wheat, rice, maize, and soybean by 6.0, 3.2, 7.4, and 3.1%, respectively (Zhao et al. 2017). In contrast, among the 12 yield regression models, only the Seonpung agronomic model reached overall significance, whereas PN showed significant model-level associations in 7 of 12 models, particularly in Seonpung and the combined dataset. These results indicate that the measured environmental and agronomic factors were more consistently associated with variation in specific yield-related traits than with final yield,

283 which may reflect the integration of multiple developmental processes and compensatory relationships
284 among yield components.
285 In Seonpung, r6d showed the strongest negative association with yield (B = -18.97), and mean yield
286 decreased from 382.6 to 264.5 kg 10 a^{-1} between the shortest (91–98 days) and longest (105–112 days) r6d
287 terciles. Previous studies have reported positive associations between seed-filling characteristics and
288 soybean yield (Egli 2004). However, r6d represents the number of days after sowing required to reach R6
289 rather than the duration or onset of seed filling, and the physiological mechanism cannot be determined
290 from the present data. In the same model, r8d showed a positive coefficient of nearly equal magnitude (B =
291 17.90); because r6d and r8d were strongly correlated (r = 0.82; Figure S2), the two coefficients should be
292 interpreted jointly, and the result is best regarded as an association between the r6d–r8d interval and yield
293 rather than an effect of r6d alone. This relationship may reflect differences in reproductive development
294 and in the seasonal environmental conditions experienced during reproductive growth, which are influenced
295 by temperature, photoperiod, and genetic factors involved in regional adaptation (Nakamichi 2015; Wang
296 et al. 2020). Analysis of stage-specific climatic conditions between R2, R4, R6, and R8 would help clarify
297 the environmental basis of this relationship.
298 Cultivar-specific ANOVA showed that the L × Y interaction for yield was larger in Seonyu2 than in
299 Seonpung (Table 3), indicating that the year-to-year change in yield varied more among locations in
300 Seonyu2. The significant interactions involving cultivar in the combined analysis likewise indicate
301 genotype × environment (G × E), an important source of soybean yield variation in multi-environment trials
302 that can complicate the identification of broadly adapted genotypes (Abebe et al. 2024). Differences in
303 flowering and maturity time may also alter the environmental conditions encountered during reproductive
304 development (Nakamichi 2015; Wang et al. 2020). Because only one medium-late-maturing and one
305 medium-maturing cultivar were examined, however, the stronger L × Y interaction in Seonyu2 should be
306 regarded as a cultivar-specific response under the environments evaluated here rather than a general
307 characteristic of shorter-duration maturity groups.
308 Among the yield-related traits, PN showed the clearest associations with the measured climatic, soil,
309 and agronomic factors. In Seonpung, the soil model was significant, with P_2O_5 positively associated with
310 PN, while rainfall showed a positive association with PN in the climate model for Seonpung and the
311 combined dataset. The positive association between rainfall and PN is consistent with the sensitivity of
312 reproductive-stage soybean to water deficit, which reduced yield by approximately 47–53% when imposed
313 at R6 (Dogan et al. 2007). Because climatic and soil predictors were recorded once per location and year,
314 and soil P_2O_5 differed markedly between years (Table 2), these associations may partly reflect year-level
315 differences and should be interpreted cautiously; neither remained significant after Holm correction (Table
316 S6). Within the Seonpung agronomic model, PN increased with NN but decreased with PD and PH. Pod

formation is a major determinant of SN and consequently an important component of soybean yield (Vogel et al. 2021). Soybean also exhibits substantial compensatory plasticity in response to plant population, with changes in branching and reproductive structures per plant contributing to yield maintenance across different densities (De Bruin and Pedersen 2009; Cox et al. 2010). The negative association between PD and PN per plant observed in Seonpung is therefore consistent with this compensatory response. However, PD in the present study was calculated from planting configuration rather than final stand counts, and the observed associations should therefore be interpreted with this limitation in mind.

BN showed significant overall models less consistently than PN, whereas only one of the 12 HSW models reached overall significance. HSW also showed substantially lower variability than PN and BN (CV = 11.9% vs. 61.4% and 55.3%; Table 1). These results suggest that the measured environmental and agronomic factors were more consistently associated with pod and branch formation than with seed weight in the present dataset. However, the limited number of significant HSW models should not be interpreted as direct evidence of stronger genetic control, because the narrow range of HSW values, statistical power, sample size, environmental range, and predictor selection also influence model significance. Notably, none of the PN, BN, or HSW models reached overall significance in Seonyu2, which may reflect both cultivar-specific responses and the smaller sample size (n = 18–20) available for this cultivar.

The weak correspondence between theoretical and observed yield may partly explain why associations observed for plant-level yield-related traits did not translate directly to final plot yield. Theoretical yield calculated from SN, HSW, and PD was not significantly correlated with observed yield ($r = 0.20$, $p = 0.21$), and the theoretical-to-observed yield ratio ranged from 0.54 to 6.42. Seed number and HSW were derived from sampled plants, whereas observed yield represented harvested plot performance, and PD was calculated from nominal planting configuration. Consequently, within-plot plant-to-plant variation, differences between nominal and realized stand density and harvest-related losses may have contributed to the discrepancy. This finding highlights the importance of matching the spatial scale of yield-component measurements with that of harvested yield when relationships between plant-level traits and plot-level productivity are evaluated.

Several limitations should be considered when interpreting these findings. First, only two cultivars with different maturity characteristics were evaluated, preventing separation of cultivar effects from broader maturity-group effects. Second, the experiment covered only two growing years, limiting the generalization of interannual responses. Third, some L × Y × C combinations contained missing or unequal numbers of observations, and climatic variables lacked within-plot replication, restricting estimation of some interaction effects. Seonyu2 lacked 2025 yield data at three locations (Goesan, Naju, and Paju), so its L × Y analysis was restricted to the five locations with data in both years. Climate and soil predictors were measured once per location and year, so replicate-level models may overstate the effective sample size. The

regression analyses were also based on relatively modest sample sizes and multiple predictor sets and should therefore be regarded primarily as assessments of association requiring validation in additional environments. Although the 48 models were not corrected for multiple comparisons across the full family, 15 of 48 (31%) reached significance, exceeding the ~5% expected by chance. Furthermore, potentially important sources of yield variation, including pest pressure, realized stand density, and detailed soil physical properties, were not included. Pod-sucking insects such as *Riptortus pedestris* and *Halyomorpha halys* can affect soybean pods and seeds in Korea (Son et al. 2000; Rahman and Lim 2017; Park et al. 2023), and such unmeasured factors may have contributed to variation not captured by the present models.

Overall, Seonpung and Seonyu2 exhibited distinct response patterns across locations and years. In Seonpung, earlier attainment of r6d, together with a longer r6d–r8d interval, was associated with higher yield, while PN showed comparatively consistent associations with climatic, soil, and agronomic factors. In Seonyu2, the larger L × Y interaction and the absence of significant regression models suggest that variation in yield and yield-related traits was not adequately represented by the predictor sets examined here. Rather than establishing r6d or PN as direct selection criteria, the present results identify them as candidate traits for further validation. Evaluation across additional cultivars, years, and environments, together with improved correspondence between plant-level sampling and plot-level measurements, will be necessary to determine their practical value for soybean cultivar evaluation and yield prediction.

References

Abebe AT, Adewumi AS, Adebayo MA, Shaahu A, Mushoriwa H, Alabi T, Derera J, Agbona A, Chigeza G. 2024. Genotype × environment interaction and yield stability of soybean (*Glycine max* L.) genotypes in multi-environment trials (METs) in Nigeria. Heliyon 10(19):e38097. https://doi.org/10.1016/j.heliyon.2024.e38097.

Anthony P, Malzer G, Sparrow S, Zhang M. 2012. Soybean yield and quality in relation to soil properties. Agron J 104(5):1443-1458. https://doi.org/10.2134/agronj2012.0095.

Cohen J. 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates, Hillsdale, NJ, USA.

Cox WJ, Cherney JH, Shields E. 2010. Soybeans compensate at low seeding rates but not at high thinning rates. Agron J 102(4):1238-1243. https://doi.org/10.2134/agronj2010.0047.

Cox WJ, Cherney JH. 2011. Growth and yield responses of soybean to row spacing and seeding rate. Agron J 103(1):123-128. https://doi.org/10.2134/agronj2010.0316.

De Bruin JL, Pedersen P. 2009. New and old soybean cultivar responses to plant density and intercepted light. Crop Sci 49(6):2225-2232. https://doi.org/10.2135/cropsci2009.02.0063.

Dogan E, Kirnak H, Copur O. 2007. Deficit irrigations during soybean reproductive stages and CROPGRO-soybean simulations under semi-arid climatic conditions. Field Crops Res 103(2):154-159. https://doi.org/10.1016/j.fcr.2007.05.009

Egli DB. 2004. Seed-fill duration and yield of grain crops. Adv Agron 83:243-279. https://doi.org/10.1016/S0065-2113(04)83005-0.

Fehr WR, Caviness CE. 1977. Stages of soybean development. Special Report 80. Iowa State University of Science and Technology, Ames, IA, USA.

Gonçalves SL, Farias JRB, Sibaldelli RNR. 2021. Soybean production and yield in the context of global climatic changes. CABI Rev 16:011. https://doi.org/10.1079/PAVSNNR202116011.

Jeung SJ, Sung JH, Kim BS. 2019. Assessment of the impacts of climate change on climatic zones over the Korean Peninsula. Adv Meteorol 2019:5418041. https://doi.org/10.1155/2019/5418041.

Kang BK, Seo JH, Kim HT, Baek IY, Choi MS, et al. 2022. Semi-early maturing, shattering resistant, large seed, and high yield soybean cultivar, 'Seonyu2ho,' for double cropping. Korean J Breed Sci 54(4):411-420. https://doi.org/10.9787/KJBS.2022.54.4.411.

Kim HT, Ko JM, Lee BW, Yun HT, Lee YH, et al. 2017. Large seed, lodging resistant and high yield soybean cultivar 'Seonpung' for soy-paste and tofu. Korean J Breed Sci 49(2):96-102. https://doi.org/10.9787/KJBS.2017.49.2.96.

Kim YH, Kong MS, Lee EJ, Lee TG, Jung GB. 2019. Status and changes in chemical properties of upland soil from 2001 to 2017 in Korea. Korean J Environ Agric 38(3):213-218.

https://doi.org/10.5338/KJEA.2019.38.3.28.
Kong MS, Kang SS, Chae MJ, Jung HI, Sonn YG, Lee DB, Kim YH. 2015. Changes of chemical properties in upland soils in Korea. Korean J Soil Sci Fert 48(6):588-592. https://doi.org/10.7745/KJSSF.2015.48.6.588.
Korea Meteorological Administration (KMA). 2021. Climatological Normals of Korea (1991-2020). Korea Meteorological Administration, Seoul, Korea.
Korean Statistical Information Service (KOSIS). 2025. Report on the 2025 cultivated area (confirmed) and crop production survey. Statistics Korea, Daejeon, Korea.
Lesk C, Anderson W, Rigden A, Coast O, Jägermeyr J, McDermid S, Davis KF, Konar M. 2022. Compound heat and moisture extreme impacts on global crop yields under climate change. Nat Rev Earth Environ 3(12):872-889. https://doi.org/10.1038/s43017-022-00368-8.
Ministry of Agriculture, Food and Rural Affairs (MAFRA). 2025. 2025 soybean production reaches 156,000 tons; close monitoring of supply and demand to ensure market stability. Ministry of Agriculture, Food and Rural Affairs, Sejong, Korea.
Nakamichi N. 2015. Adaptation to the local environment by modifications of the photoperiod response in crops. Plant Cell Physiol 56(4):594-604. https://doi.org/10.1093/pcp/pcu181.
Nelson RL. 1986. Defining the seed-filling period in soybeans to predict yield. Crop Sci 26(1):132-135. https://doi.org/10.2135/cropsci1986.0011183X002600010031x.
Park SB, Koo HN, Seok SJ, Kim HK, Yi HJ, Kim GH. 2023. Feeding behavior comparison of bean bugs, Riptortus pedestris and Halyomorpha halys, on different soybean cultivars. Insects 14(4):322. https://doi.org/10.3390/insects14040322.
Rahman G, Kim JY, Kim TW, Park M, Kwon HH. 2025. Spatial and temporal variations in temperature and precipitation trends in South Korea over the past half-century (1974-2023) using innovative trend analysis. J Hydro-Environ Res 58:1-18. https://doi.org/10.1016/j.jher.2024.11.002.
Rahman MM, Lim UT. 2017. Evaluation of mature soybean pods as a food source for two pod-sucking bugs, *Riptortus pedestris* (Hemiptera: Alydidae) and *Halyomorpha halys* (Hemiptera: Pentatomidae). PLoS ONE 12(4):e0176187. https://doi.org/10.1371/journal.pone.0176187.
Roh AS, Park JS, Park YS, Ju OJ, Shin MW, Kang SS. 2018. Status and changes in chemical properties of upland soil in Gyeonggi Province. Korean J Soil Sci Fert 51(4):435-444. https://doi.org/10.7745/KJSSF.2018.51.4.435.
Seo J, Park J, Choi M, Jung K, Chun H, Lee S, Gong D, Chae SE, Jeon SH, Yoon DK. 2024. Growth and yield response of soybean (*Glycine max* L.) in relation to sowing date in the southern region of South Korea. Agronomy 14(11):2624. https://doi.org/10.3390/agronomy14112624.
Smith JR, Nelson RL. 1986a. Selection for seed-filling period in soybean. Crop Sci 26(3):466-469.

https://doi.org/10.2135/cropsci1986.0011183X002600030005x.
Smith JR, Nelson RL. 1986b. Relationship between seed-filling period and yield among soybean breeding lines. Crop Sci 26(3):469-472. https://doi.org/10.2135/cropsci1986.0011183X002600030006x.
Son CK, Park SG, Hwang YH, Choi BS. 2000. Field occurrence of stink bug and its damage in soybean. Korean J Crop Sci 45(6):405-410.
Vogel JT, Liu W, Olhoft P, Crafts-Brandner SJ, Pennycooke JC, Christiansen N. 2021. Soybean yield formation physiology: a foundation for precision breeding based improvement. Front Plant Sci 12:719706. https://doi.org/10.3389/fpls.2021.719706.
Wang L, Sun S, Wu T, Liu L, Sun X, et al. 2020. Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. Plant Biotechnol J 18(9):1869-1881. https://doi.org/10.1111/pbi.13346.
Yoon BS, Choi SC, Lim SJ, Heo SJ, Seo Y. 2020. Status and changes in chemical properties of upland soil in Gangwon Province between 2001 and 2017. Korean J Soil Sci Fert 53(4):549-557. https://doi.org/10.7745/KJSSF.2020.53.4.549.
Zhao C, Liu B, Piao S, et al. 2017. Temperature increase reduces global yields of major crops in four independent estimates. Proc Natl Acad Sci USA 114(35):9326-9331. https://doi.org/10.1073/pnas.17017621.

Table 1. Agronomic and phenological traits of two soybean cultivars across eight locations and two years (2024–2025).

Trait	Seonpung	Seonyu2	Overall
PH (cm)	59.5 ± 11.9 (20.0)	43.8 ± 14.1 (32.3)	52.7 ± 15.1 (28.6)
FPH (cm)	17.2 ± 8.3 (48.4)	11.0 ± 6.3 (56.6)	14.7 ± 8.1 (55.5)
NN	15.6 ± 2.8 (18.1)	12.3 ± 2.0 (16.4)	14.2 ± 3.0 (20.9)
BN	3.0 ± 1.8 (58.2)	2.9 ± 1.5 (50.4)	3.0 ± 1.6 (55.3)
PN	54.7 ± 34.7 (63.5)	59.5 ± 34.7 (58.4)	56.7 ± 34.8 (61.4)
SN	113.8 ± 84.1 (74.0)	116.7 ± 77.3 (66.2)	115.0 ± 81.3 (70.7)
SPP	2.0 ± 0.5 (26.5)	1.9 ± 0.4 (19.2)	2.0 ± 0.5 (24.1)
r2d (DAS)	58.8 ± 5.6 (9.6)	48.5 ± 7.7 (15.9)	54.2 ± 8.4 (15.4)
r4d (DAS)	75.8 ± 8.2 (10.8)	62.7 ± 7.5 (12.0)	70.0 ± 10.2 (14.6)
r6d (DAS)	101.3 ± 6.9 (6.8)	92.7 ± 6.9 (7.4)	97.4 ± 8.1 (8.3)
r8d (DAS)	132.8 ± 7.1 (5.4)	115.5 ± 12.0 (10.4)	125.1 ± 12.9 (10.3)
HSW (g)	24.5 ± 2.7 (11.2)	23.2 ± 2.8 (12.2)	23.9 ± 2.8 (11.9)
Yield (kg $10a^{-1}$)	315.6 ± 94.8 (30.0)	230.6 ± 104.0 (45.1)	280.8 ± 106.5 (37.9)

Values are presented as mean ± SD (CV, %). PH, plant height; FPH, first-pod height; NN, node number; BN, branch number; PN, pod number; SN, seed number; SPP, seeds per pod; HSW, hundred-seed weight. r2d, r4d, r6d, and r8d indicate days after sowing (DAS) to the R2, R4, R6, and R8 growth stages, respectively. Trait values by location, year, and cultivar are provided in Table S3.

462 Table 2. Climatic and soil properties across eight experimental locations in 2024 and 2025.

Variable	2024	2025	Overall
Tmin (°C)	7.61 ± 1.78 (23.4)	8.36 ± 3.20 (38.2)	7.98 ± 2.53 (31.7)
Tmax (°C)	35.83 ± 0.39 (1.1)	38.66 ± 1.70 (4.4)	37.24 ± 1.89 (5.1)
Tmean (°C)	24.39 ± 0.85 (3.5)	24.51 ± 0.54 (2.2)	24.45 ± 0.69 (2.8)
Rainfall (mm)	939.1 ± 269.6 (28.7)	1013.8 ± 257.5 (25.4)	976.5 ± 257.6 (26.4)
Sunshine duration (h)	785.7 ± 88.3 (11.2)	760.2 ± 135.1 (17.8)	773.0 ± 111.0 (14.4)
Solar radiation (MJ m^{-2})	2145.8 ± 243.2 (11.3)	2237.6 ± 291.7 (13.0)	2191.7 ± 263.7 (12.0)
pH (1:5)	6.75 ± 0.66 (9.7)	6.67 ± 0.69 (10.3)	6.71 ± 0.65 (9.7)
OM (%)	17.07 ± 1.60 (9.4)	2.24 ± 0.99 (44.4)	9.65 ± 7.77 (80.5)
TN (% N)	0.08 ± 0.03 (35.7)	0.14 ± 0.04 (25.1)	0.11 ± 0.04 (39.7)
P_2O_5 (mg kg^{-1})	14.65 ± 5.11 (34.9)	426.34 ± 353.78 (83.0)	220.49 ± 321.90 (146.0)
EC (dS m^{-1})	0.14 ± 0.03 (22.1)	0.64 ± 0.53 (82.9)	0.39 ± 0.45 (114.8)

463 Values are presented as mean ± SD (CV, %) across the eight locations in each year. Tmin, Tmax, and
464 Tmean, minimum, maximum, and mean air temperature during the growing season; rainfall, sunshine
465 duration, and solar radiation are cumulative totals over the growing season. EC, electrical conductivity; TN,
466 total nitrogen; OM, organic matter; P_2O_5, available phosphate. Climatic and soil data by location and year
467 are provided in Supplementary Table S4.
468

Table 3. Analysis of variance for agronomic traits, phenological development, and yield across locations, years and cultivars.

Analysis	Source	PH	FPH	NN	BN	PN	SN	SPP	r2d	r4d	r6d	r8d	HSW	Yield
Combined	L	115.75 ***	15.31 ***	19.18 ***	21.71 ***	27.72 ***	47.89 ***	13.29 ***	902.04 ***	3794.94 ***	846.29 ***	117.37 ***	2.32 ns	13.11 ***
	Y	283.30 ***	77.00 ***	33.77 ***	19.06 ***	62.79 ***	78.07 ***	7.38 **	1160.33 ***	208.33 ***	0.07 ns	110.82 ***	0.54 ns	10.87 **
	C	1161.05 ***	147.66 ***	659.85 ***	0.42 ns	0.52 ns	2.59 ns	25.56 ***	3675.00 ***	27840.33 ***	2640.07 ***	768.05 ***	0.99 ns	37.27 ***
	L × Y	42.90 ***	10.02 ***	7.84 ***	4.02 ***	8.42 ***	9.25 ***	1.82 ns	459.01 ***	2105.69 ***	241.28 ***	43.73 ***	3.31 *	23.17 ***
	L × C	26.30 ***	11.93 ***	2.75 *	6.27 ***	6.93 ***	10.30 ***	1.88 ns	280.43 ***	2401.50 ***	29.22 ***	34.53 ***	2.00 ns	4.51 **
	Y × C	13.65 ***	0.10 ns	9.59 **	0.81 ns	0.49 ns	0.02 ns	0.41 ns	507.00 ***	1408.33 ***	232.07 ***	50.26 ***	8.07 **	0.10 ns
	L × Y × C	11.39 ***	1.48 ns	7.12 ***	8.65 ***	20.20 ***	25.35 ***	0.86 ns	251.02 ***	1533.16 ***	266.98 ***	16.09 ***	4.13 **	7.51 ***
Seonpung	L	55.44 ***	12.21 ***	11.78 ***	19.86 ***	24.54 ***	43.68 ***	13.00 ***	6.10 **	3.05 *	6.74 ***	6.13 **	12.95 ***	7.06 ***
	Y	60.52 ***	23.35 ***	1.56 ns	2.13 ns	24.22 ***	45.70 ***	0.92 ns	0.05 ns	3.05 ns	9.08 **	0.09 ns	35.84 ***	15.94 ***
	L × Y	12.66 ***	4.25 ***	2.45 *	5.68 ***	8.98 ***	11.45 ***	0.94 ns	NE	NE	NE	NE	13.35 ***	5.83 ***
Seonyu2	L	107.81 ***	13.11 ***	14.83 ***	6.42 ***	8.78 ***	13.51 ***	8.65 ***	303.15 ***	1362.78 ***	131.13 ***	43.39 ***	1.49 ns	13.48 ***
	Y	361.74 ***	58.15 ***	56.96 ***	17.22 ***	34.29 ***	36.41 ***	10.45 **	631.84 ***	532.89 ***	44.26 **	61.25 ***	1.29 ns	9.91 *
	L × Y	73.43 ***	9.07 ***	15.86 ***	9.15 ***	18.84 ***	23.03 ***	3.75 **	NE	NE	NE	NE	3.96 *	61.20 ***

Values are *F*-values from Type III ANOVA. L, location; Y, year; C, cultivar. ns, *, **, and *** indicate non-significance or significance at $p <$ 0.05, 0.01, and 0.001, respectively; NE, not estimable.

Table 4. Significant multiple regression models for yield and yield-related traits by cultivar and predictor category.

Cultivar	Category	DV	R^2	Adj. R^2	F	p	Predictors (B, p)
Seonpung	Agronomic	Yield	0.66	0.51	4.37	0.007	r6d (-18.97, p<0.001); r8d (17.90, p=0.001); FPH (-9.63, p=0.039); NN (-23.78, p=0.084)
		PN	0.56	0.44	4.55	0.007	PD (-0.00, p<0.001); NN (11.95, p=0.005); PH (-1.03, p=0.040)
		BN	0.54	0.42	4.29	0.010	NN (0.50, p=0.014); PH (-0.06, p=0.021); r6d (-0.07, p=0.043); PD (-0.00, p=0.080)
	Climate	PN	0.54	0.45	5.62	0.004	Rainfall (0.07, p=0.001); Tmin (4.91, p=0.048)
		BN	0.49	0.38	4.53	0.010	Sunshine duration (-0.01, p=0.007); Rainfall (0.00, p=0.042)
	Soil	PN	0.45	0.40	8.74	0.002	P_2O_5 (0.06, p=0.001)
	Climate + Soil	PN	0.55	0.43	4.40	0.009	Rainfall (0.08, p=0.003); Tmin (5.18, p=0.046)
		BN	0.54	0.41	4.16	0.011	Sunshine duration (-0.01, p=0.005); Rainfall (0.00, p=0.018)
Combined	Agronomic	BN	0.23	0.15	2.77	0.042	r6d (-0.04, p=0.079); HSW (0.12, p=0.097)
	Climate	PN	0.43	0.36	5.80	<0.001	Rainfall (0.06, p<0.001); Tmax (4.03, p=0.098)
		BN	0.31	0.22	3.36	0.013	Rainfall (0.00, p=0.038)
	Soil	PN	0.21	0.15	3.46	0.025	P_2O_5 (0.04, p=0.034)
	Climate + Soil	PN	0.45	0.34	4.20	0.002	Rainfall (0.06, p=0.001)
		BN	0.32	0.19	2.46	0.036	Rainfall (0.00, p=0.055); Sunshine duration (-0.00, p=0.067)
		HSW	0.37	0.24	2.88	0.017	Tmean (-1.92, p=0.001); pH (1.35, p=0.046)

Only statistically significant models ($p < 0.05$) are shown. Predictors with $p < 0.10$ are listed with their unstandardized regression coefficients (B) and unadjusted p values; Holm-adjusted p values are given in Table S6. DV, dependent variable; B, unstandardized regression coefficient; adj. R^2, adjusted R^2.

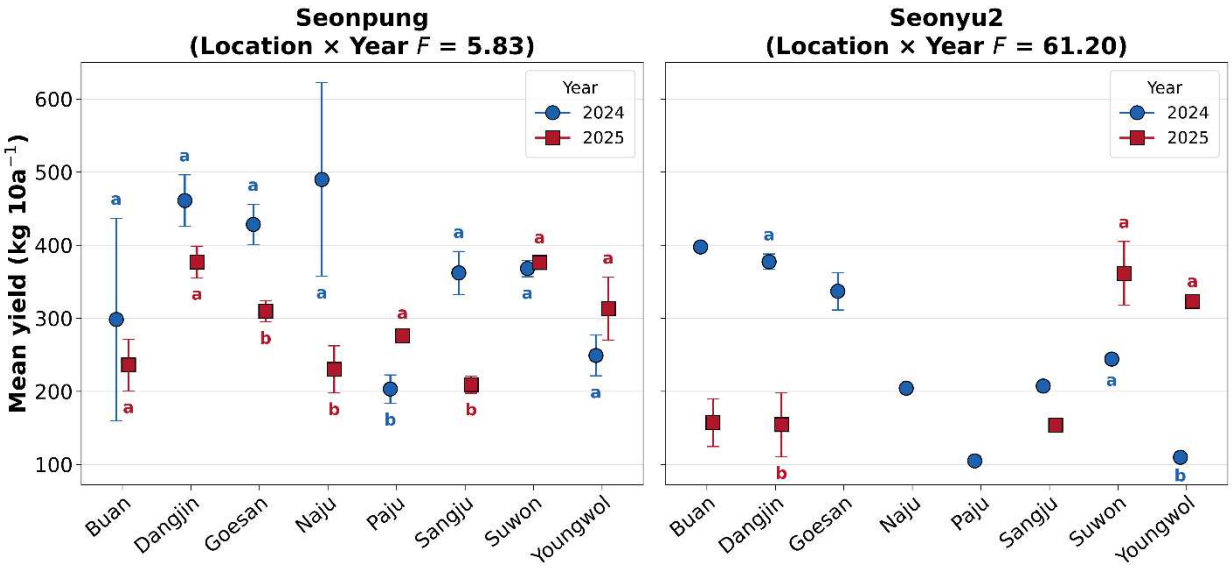


Figure 1. Yield variation across locations and years by cultivar. Points show location × year means ± SD from field replicates. Letters indicate Tukey–Kramer HSD comparisons of 2024 vs. 2025 within each location ($p < 0.05$). No letters are shown where a year has only one replicate. No 2025 yield data are available for Seonyu2 at Goesan, Naju, and Paju.

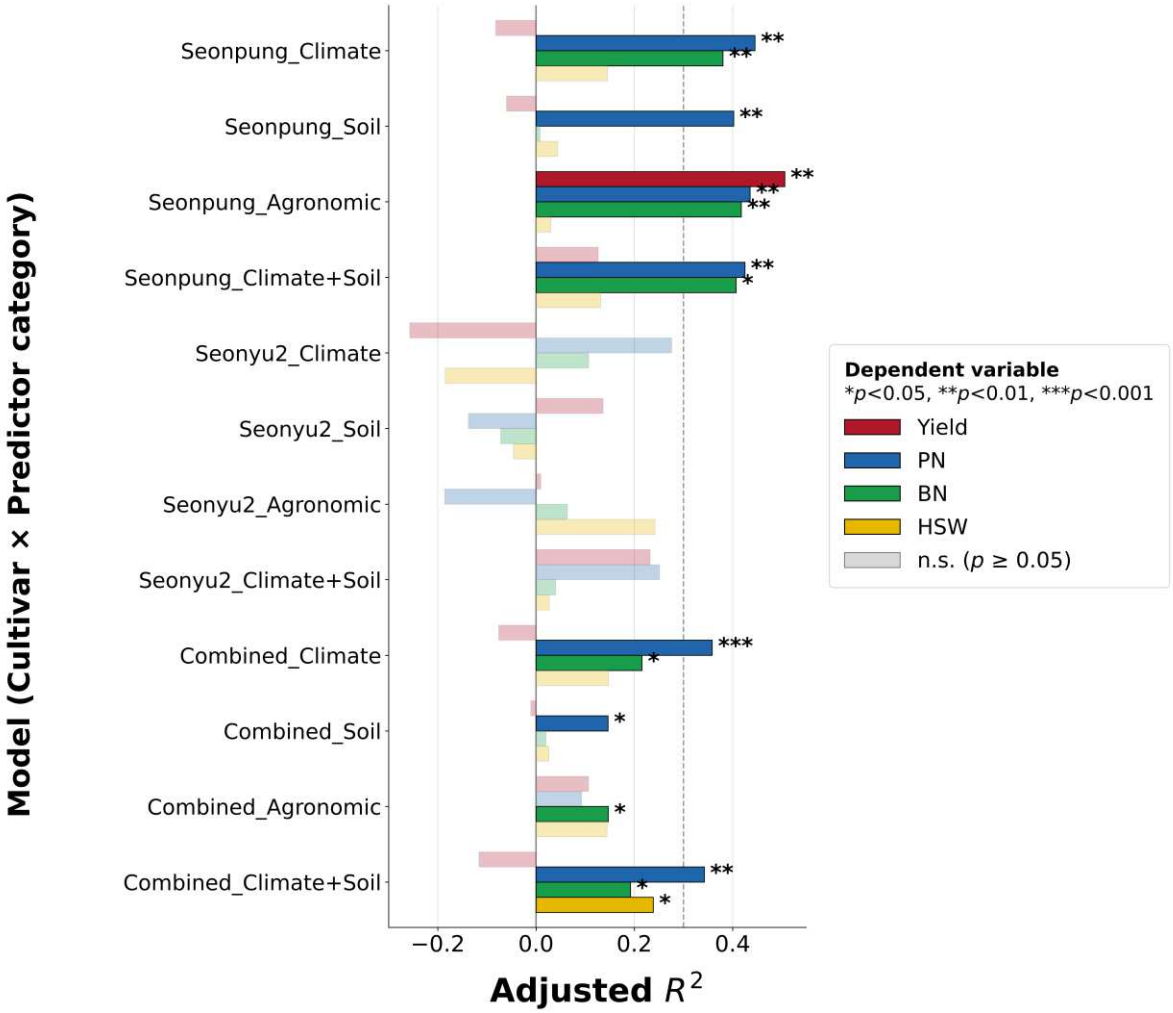


Figure 2. Adjusted R^2 for yield and three yield-related traits (pod number, branch number, hundred-seed weight) across 48 regression models (3 datasets × 4 predictor categories × 4 dependent variables). Bars are shown at full color with an asterisk where the model was statistically significant ($p < 0.05$); non-significant models are shown faded. Dashed line indicates adjusted $R^2 = 0.3$.

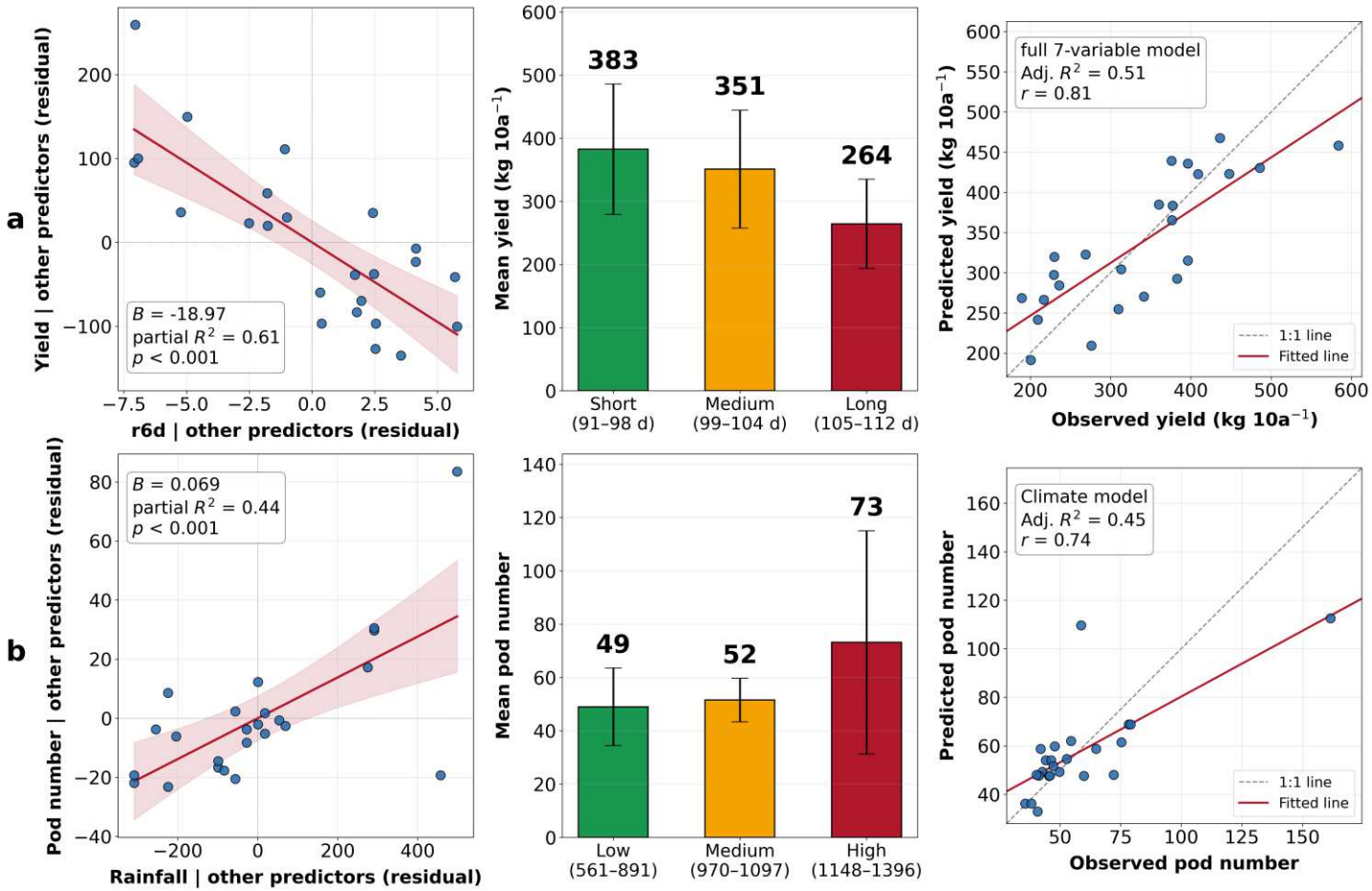


Figure 3. Associations of r6d with yield (a) and rainfall with pod number (b) in Seonpung. (a) Left to right: partial regression of yield on r6d, controlling for other predictors in the agronomic-trait model; mean yield by r6d tercile group; observed vs. predicted yield from the same model. (b) Left to right: partial regression of pod number on rainfall, controlling for other predictors in the climate model; mean pod number by rainfall tercile group; observed vs. predicted pod number from the same model. *B*, unstandardized coefficient; shaded bands, 95% confidence intervals; error bars, ± SD.

Supplementary Files

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

- Supplementaryfiguresvf1.pptx
- Supplementarytablesvf1.xlsx