Saffron Agronomy & Technology
Homepage: http://saffron.torbath.ac.ir
Vol. 13, No. 3, 2025, p. 291-306

Determining Key Agronomic and Environmental Drivers of Saffron (*Crocus sativus* L.) Yield and Quality Using LASSO Regression

Abdollah Dorpoor Sorkhsarayi[1], Mehdi Rastgoo[2*], Ghorbanali Asadi[3], Javad Feizy[4] and Mohammad Hassan Fallah[5]

Article type:
Research Article
Article history:
Submitted 20 November 2025
Revised: 13 December 2025
Accepted: 28 April 2026
Available Online: 30 April 2026

How to cite this article:
Dorpoor Sorkhsarayi, A., Rastgoo, M., Asadi, Gh., Feizy, J., and Fallah, M. H. (2025). Determining Key Agronomic and Environmental Drivers of Saffron (*Crocus sativus* L.) Yield and Quality Using LASSO Regression, *Saffron Agronomy & Technology, 13*(3), 291-306. https://doi.org/10.22048/jsat.2026.558332.1575

Abstract

Saffron is a high-value crop of strategic and economic importance in Iran, particularly in semi-arid regions. Its productivity and quality are influenced by complex interactions among agronomic, environmental, and management factors. This study aimed to identify and analyze the key determinants of saffron yield (stigma yield per hectare and per kilogram of fresh flowers) and quality indices (crocin, picrocrocin, and safranal) using a data-driven approach. In 2023, data were collected from 99 saffron farms across eight counties in Razavi Khorasan Province, encompassing 75 variables that recorded climate, soil, management practices, and farmer demographics. Least Absolute Shrinkage and Selection Operator (LASSO) regression with 10-fold cross-validation was applied for variable selection and predictive analysis. Results indicated that stigma yield per hectare was primarily influenced by corm planting rate, organic fertilizer, corm weight, and field area, achieving $R^2 = 0.63$ and RMSE = 3.75 kg.ha^{-1}. For stigma yield per kilogram of fresh flowers, phosphorus fertilization, corm weight, and planting density were the strongest positive predictors, with $R^2 = 0.70$ and RMSE = 0.69 g.kg^{-1}. Moderate positive effects were observed for organic fertilizer and irrigation frequency, while quadratic effects suggested threshold responses for corm size and irrigation. For quality traits, phosphorus was the dominant positive predictor of crocin ($\beta = 18.3$) and picrocrocin ($\beta = 3.97$), whereas altitude and foliar spray frequency negatively affected picrocrocin and safranal. The effects of nitrogen and sulfur fertilizers were minor and nonlinear. Simplified models retained predictive accuracy ($R^2 = 0.70$), improving practical applicability. These findings highlight the importance of site-specific phosphorus management, corm quality monitoring, and optimized irrigation for enhancing saffron yield and quality. LASSO regression effectively identified influential variables, supporting precision agriculture and decision-support tools for sustainable saffron production under semi-arid conditions.

Keywords: Climatic factors, Irrigation scheduling, Nonlinear effects, Nutrient management, Picrocrocin, Variable selection

1 - PhD student in Agrotechnology, Weed Science, Faculty of Agriculture, Ferdowsi University of Mashhad
2 - Professor, Department of Agrotechnology, Faculty of Agriculture, Ferdowsi University of Mashhad
3 - Professor, Department of Agrotechnology, Faculty of Agriculture, Ferdowsi University of Mashhad
4 - Associate Professor, Department of Food Safety and Quality Control, Research Institute of Food Science and Technology (RIFST), Mashhad, Iran
5 - PhD graduate in Crop Ecology, Department of Agrotechnology, Faculty of Agriculture, Ferdowsi University of Mashhad
Corresponding author email: m.rastgoo@um.ac.ir
https://doi.org/10.22048/jsat.2026.558332.1575

Introduction

Saffron (*Crocus sativus* L.) is widely recognized as one of the world's prized medicinal and spice crops, thanks to its unique bioactive properties and economic benefits. The excellent adaptability of saffron to arid and semi-arid climates and its low water requirement make it well-suited for cultivation in areas such as Iran (Mohammadkhani et al., 2023). While saffron is commonly referred to as "red gold," not only is it high in market value, but the red gold contribution is large in rural economic support by providing jobs and a sustainable income opportunity for rural livelihoods, particularly in areas with low or limited water supply (Riyahi et al., 2023). As the world's largest producer and exporter of saffron, Iran mainly produces in the Khorasan Razavi Province, with further growth potential. Despite being the largest producer, saffron yields in Iran are lower than their genetic potential, suggesting that widespread gains in production could be achieved with appropriate agronomic and management practices (Pirasteh-Anosheh et al., 2023).

Saffron is primarily propagated using corms, and the yield and quality of its stigmas with economic value will depend on many biological, environmental, and agronomic considerations. The agronomic considerations that will affect saffron productivity include corm quality (weight, viability, etc.), corm density and depth of planting, fertilization (organic versus 'chemical'), and irrigation timing (Azizi et al., 2023). This is exemplified by healthy corms of the correct weight, resulting in a greater number of flowers and better-quality stigma, while the correct depth of planting helps protect the corms from extreme temperatures and moisture (Ziaei et al., 2023).

In addition to these agronomic factors, the socio-economic characteristics of farmers, such as age, education, experience, and access to agricultural extension services, influence management decisions and strawberry-related outcomes. These characteristics lead to different ways they adopt best practices and influence their saffron productivity (Shahnoushi et al., 2020). Water resource management under drought stress underscores the importance of efficient water management during saffron cultivation. Factors such as the timing of the first irrigation and the number of irrigation events strongly influence plant development and flower production. Irrigation is also important because it offers dual benefits: it improves saffron yield and conserves limited water resources (Sepaskhah & Kamkar-Haghighi, 2009).

Fertilizer management is also key to saffron performance. A balanced application of organic fertilizers improves soil physical and chemical properties, increases moisture retention, and enhances nutrient availability, especially when the soil is nutrient-poor. At the same time, the proper use of inorganic macronutrients such as nitrogen, phosphorus, and potassium supports vegetative growth and high-quality flowering. The fertilizer strategy must be guided by the soil and the crop's needs. This is a vital step in maximizing yield and quality (Wang et al., 2013).

The climatic conditions can affect saffron productivity and quality. Saffron is typically grown in areas with high solar radiation and moderate temperatures, with a favorable growing temperature range of 15 to 20°C. Deviation from the temperature range can disrupt physiological functions such as photosynthesis, flowering, and stigmatic development (Pirasteh-Anosheh et al., 2023; Wang et al., 2021). Additionally, humidity and the timing of rainfall or moisture vary and affect disease incidence and crop health; excessive moisture during flowering can lead to fungal infections and lower stigma quality, whereas no moisture during the critical growth stage can improve flowering and photosynthesis (Gresta et al., 2009). Therefore, understanding the interactions between climatic factors and saffron growth can help producers develop appropriate growth practices for their particular agroecological zone.

As saffron is a significant cash crop for the Iranian upstream agricultural economy and rural livelihoods, it is important to evaluate the agronomic, managerial, and environmental factors that impact yield and quality across all relevant saffron production systems. This study examined relevant agronomic, managerial, and environmental factors in saffron production across farms in Khorasan Razavi Province and considered the changing socio-economic characteristics of farmers, as well as how planting methods, nutrient management, irrigation practices, and climatic conditions influence saffron yield and quality. Unlike conventional agronomic studies that typically focus on a limited number of variables under controlled conditions, this study adopts a high-dimensional, data-driven approach to simultaneously evaluate a wide range of agronomic, environmental, and socio-economic factors. This enables the identification of dominant drivers and their relative importance within a complex production system. The findings of this research are intended to suggest improvements in saffron management systems and to enhance the yield and quality of saffron production in this important saffron cultivation area of the country.

Materials and Methods

Study region and sampling framework

In 2023, the research was conducted across Razavi Khorasan Province in northeastern Iran. This area is a global center of saffron production, accounting for over 85% of global output (Koocheki & Sabet Teimouri, 2014). The saffron-producing counties of Razavi Khorasan vary in climate, soil physical and chemical properties, and agronomic practices across the landscape. The sample size was not determined using a predefined statistical formula; rather, it was based on achieving adequate representation of saffron production systems across the study region. A stratified proportional sampling approach was used, where the number of farms selected from each county was proportional to its cultivation area. The final sample size (n = 99) was determined based on regional coverage, variability in management practices, and logistical feasibility. This approach is commonly used in observational agricultural studies where the goal is to capture system variability rather than estimate population means. Accordingly, we intentionally selected eight major counties based on the area of saffron production: Zaveh (27 samples), Torbat-e Heydarieh (14 samples), Kashmar and Khalilabad (6 samples), Roshtkhar and Khaf (5 samples), Neyshabur (12 samples), Taybad (8 samples), Bakharz (9 samples), and Torbat-e Jam (18 samples). Together, these counties account for more than 60% of the province's total saffron area, and each farm was assigned GPS coordinates for spatial representation (Figure 1).

We used a stratified sampling strategy to ensure proportional representation of saffron cultivation areas across counties, which is a commonly recommended approach for capturing spatial and management variability in agricultural systems (e.g., Cochran, 1977).

Farms were selected based on a history of regular saffron production, including varying levels of input intensity (e.g., high- or low-input systems). This design provided a comprehensive assessment of climatic and farm-management variability affecting saffron productivity.

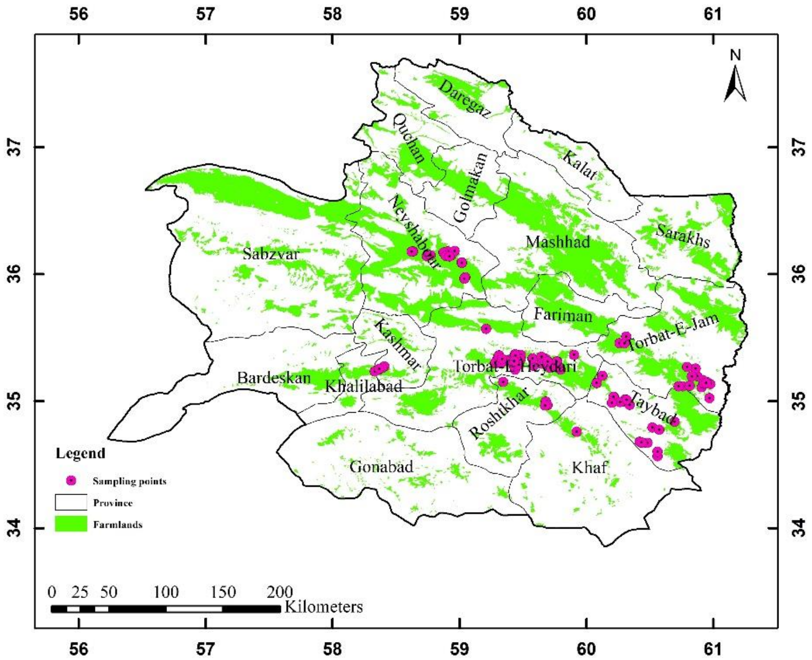

Figure 1. Geographical location of the study area and spatial distribution of sampled saffron farms in Razavi Khorasan Province.

Agronomic and management data collection

Data on agronomic and management details were derived via face-to-face interviews using a structured questionnaire administered by trained agricultural extension agents. The questionnaire contained both qualitative and quantitative variables and included: farmer demographics: age, education level; farm characteristics: area of planting, age of field (years since planted initially); characteristics of planted corms (e.g., average weight per hectare), depth of planting; nutrient management variables: including application rates of organic amendments (manure or compost) and chemical fertilizers (N, P, and K); irrigation management indicators: including irrigation frequency (number of events) and timing of the first irrigation; irrigation volume was not available and therefore not included; irrigation regimes: number of irrigation events and the timing of first irrigation; and the number of foliar sprays (micronutrients). These variables have been indicated to have substantial effects on saffron yield and quality (Koocheki & Sabet Teimouri, 2014). Also, to quantify stigma quality indices, 1 kg of fresh saffron flowers was collected from each sampled farm at the full-flowering stage. Sampling was conducted across all aspects of true randomness, with several random samples collected from different areas of the farm field to maximize spatial representativeness. Flower samples were collected at harvest, and stigmas were separated and dried under standardized conditions. All samples were processed within 2 hours after harvest to minimize variability due to post-harvest changes.

Soil sampling and analysis

Soil samples were collected from 0-30 cm depth at each site using a soil auger, following a zigzag sampling design. Each farm was considered as a single observational unit in the analysis. From each farm, 5 samples were collected to account for within-field variability. The measurements obtained from these subsamples were averaged to produce a single representative value per farm for each variable before statistical analysis. Samples were air-dried, sieved (2 mm), and sent to a certified soil laboratory for analysis. The following parameters were analyzed: pH, electrical conductivity (EC; $dS.m^{-1}$), organic matter content (OM), total nitrogen (%), available phosphorus ($mg.kg^{-1}$), extractable potassium ($mg.kg^{-1}$), and soil texture fractions (clay, silt, sand). Soil samples were analyzed following standard laboratory protocol. Soil texture analysis was performed by the Bouyoucos hydrometer method (Bouyoucos, 1936). Soil pH was measured in a 1:2.5 soil-water suspension with a glass-calomel electrode (MP 220 AFAB Lab, LLC) (Muche et al., 2015). Organic matter was estimated using the Walkley-Black method (Walkley and Black,1934), while total nitrogen was determined using Kjeldahl digestion (Bradstreet, 1954). Available phosphorus was extracted using the Olsen method (Olsen, 1954). K was determined by 1 N ammonium acetate extraction and measured by flame photometry using the same extract (Muche et al., 2015). Soil characteristics are important determinants of nutrient availability and corm development in saffron systems (Cardone et al., 2020).

Climatic data collection

Climatic data were collected from the Khorasan Razavi Meteorological Organization's database within and around the study area. To estimate climatic variables at each farm location, a K-nearest neighbors (KNN) spatial interpolation method was applied. This approach uses data from the nearest meteorological stations to estimate local climatic conditions for each farm, thereby improving spatial representation across regions, including areas without direct station coverage. Monthly and yearly averages of minimum (Tmin), maximum (Tmax), and mean (Tmean) temperatures, as well as total monthly and yearly precipitation for saffron cultivation from October 2022 to September 2023 to cover both the active growth period and dormancy phase of saffron, were collected for each province. Temperature and precipitation are key climatic drivers of saffron yield, influencing flowering induction, growth dynamics, and overall productivity (Pirasteh-Anosheh et al., 2023).

Assessment of saffron quality indices

The saffron quality indices were determined at the Saffron Institute Laboratory of the Mashhad Research Institute of Food Science and Technology. Crocin, picrocrocin, and safranal contents were measured using UV-Visible spectrophotometry according to the ISO/TS 3632 standard. For each sample, 500 mg of ground saffron was mixed with 900 mL of distilled water in a 1000 mL volumetric flask and stirred at 1000 rpm for 1 hour. The solution was then adjusted to 1000 ml, homogenized, and then diluted (20 ml to 200 ml). After filtering in low-light conditions, the absorbance (A) was recorded with distilled water used as a reference between 200–700 nm. Absorbance at 257 nm indicated picrocrocin, 330 nm indicated safranal, and 440 nm indicated crocin. The concentration (E) of each molecule was calculated using equation 1:

$$E = \frac{D \times 1000}{m(100 - H)} \qquad \text{(Eq.1)}$$

Where D = absorbance, m = sample weight (g), and H = moisture and volatile substances content (%) (Kaveh & Salari, 2018). The three bioactive substances in saffron, crocin, picrocrocin, and safranal, determine the color, the taste, and the aroma of saffron, respectively. Previous references have accounted for the significance of bioactive substances in the grading and standardization of saffron products (Lozano et al., 2000).

Data reprocessing

The original dataset included 115 observations. After removing records with missing, inconsistent, or extreme values (identified using the interquartile range (IQR) method and Mahalanobis distance (Ghorbani, 2019), 99 valid observations remained. All continuous predictor variables were standardized (zero mean, unit variance) to eliminate scale effects and improve numerical stability in regression modeling. All variables were standardized prior to analysis; therefore, the reported regression coefficients represent standardized effects, indicating the relative importance of predictors rather than their effects in original units. This allows direct comparison of effect sizes across variables measured on different scales.

Statistical analysis and predictive modeling

The study examined five response variables: stigma yield per hectare ($kg.ha^{-1}$), stigma yield per kilogram of fresh flowers ($g.kg^{-1}$), and three quality indices, including crocin (color), picrocrocin (taste), and safranal (aroma). A total of 75 predictor variables were initially considered, grouped into three categories: (1) climatic variables (n = 54), including monthly and annual minimum, maximum, and mean temperatures (Tmin, Tmax, Tmean), precipitation, as well as latitude and altitude; (2) soil characteristics (n = 9), including electrical conductivity (EC), pH, organic matter, soil texture, and macronutrient concentrations; (3) management and farmer-related variables (n = 12), including field area, field age, characteristics of planted corms (e.g., weight and density), irrigation management indicators (frequency and timing), fertilizer application rates, and selected demographic attributes of farmers.

Surgical quadratic terms for selected continuous predictors (e.g., temperature and nutrient levels) were added to account for non-linear effects. To account for potential nonlinear relationships, quadratic terms were included for selected continuous variables. This approach allows the detection of threshold or optimum responses, which are common in agronomic systems.

All predictor variables were standardized (mean = 0, standard deviation = 1) prior to model fitting, while response variables were kept in their original scale. Quadratic terms were generated from the standardized variables to ensure consistency and avoid scaling bias. For example, both insufficient and excessive phosphorus application may negatively affect plant performance, indicating a nonlinear response.

The LASSO regression method (Least Absolute Shrinkage and Selection Operator) was used to select the most predictive variables and exclude correlated variables. LASSO is a penalized regression method that performs covariate selection and regularization, resulting in sparse models that are easy to interpret (Tibshirani, 1996). LASSO computes the optimal penalty parameter (λ), using 10-fold cross-validation, which minimizes the mean cross-validated error. The LASSO objective function is presented in equation 2:

$$\min_{\beta_0,\beta}\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda\sum_{j=1}^{p}\left|\beta_j\right|\right\} \quad \text{(Eq. 2)}$$

where yi denotes the response variable for the i-th observation (including yield and quality traits), xij is the value of the j-th predictor variable for the i-th observation (from the set of 75 climatic, soil, and management variables), βj are the corresponding regression coefficients, and λ is the regularization parameter controlling the degree of coefficient shrinkage. The optimal value of λ was selected using 10-fold cross-validation. For quantitative response variables, the λ_{1se} criterion was applied to obtain simpler and more stable models, while for quality-related traits, λ_min was used to maximize predictive accuracy. The selected λ values for each model are reported in Tables 1 and 4. The λ_{1se} criterion was preferred for yield models to reduce overfitting and improve model interpretability.

The final predictive model was developed using the LASSO-selected variables with nonzero coefficients. Model performance was evaluated using MSE, RMSE, R^2, adjusted R^2, SE, and the index of agreement (d) (Willmott, 1981), which measures the agreement between observed and predicted values. These metrics facilitated the identification of key factors influencing saffron yield and quality. This analysis enabled investigations of key agronomic, environmental, and management factors influencing saffron yield and quality, with practical recommendations for improved cultivation and production practices.

Results and Discussion

Saffron stigma yield

LASSO regression models were applied to

predict two key saffron yield traits: stigma yield per hectare ($kg.ha^{-1}$) and stigma yield per kilogram of fresh flowers ($g.kg^{-1}$) (Table 1). Model selection used the one-standard-error rule (λ_{1se}), balancing model simplicity and predictive accuracy. For stigma yield per hectare, the optimal λ (0.8299) produced an MSE of 14.06 (SE = 1.885) with six predictors retained. In contrast, the model for stigma yield per kilogram of fresh flowers used $\lambda = 0.132$, achieving an MSE of 0.599 (SE = 0.041) with 11 predictors. These results indicate that LASSO effectively identifies influential predictors while preventing overfitting.

Predictors were classified into "High", "Medium", and "Low" importance based on their relative contribution (Table 2). The absolute value of each standardized coefficient was divided by the maximum absolute coefficient in the model. Variables with ≥ 70% of the maximum were considered "High", 30–70% as "Medium", and < 30% as "Low". This approach allows comparison of predictor importance within each model regardless of scale differences.

For stigma yield per hectare, the most important predictors were the amount of corms used (standardized coefficient $\beta = 3.932$), and manure fertilizer ($\beta = 1.838$), indicating a strong positive association with stigma yield per hectare, while average corm weight, field area, and irrigation frequency had smaller positive effects (Table 2). For stigma yield per kilogram of fresh flowers, the most influential factors were average corm weight (0.429), phosphorus fertilizer (0.319), and corm planting weight ($\beta = 0.300$) rate (0.300), with moderate contributions from irrigation frequency, organic fertilizer, and maximum December temperature. Minor effects were observed for nitrogen fertilizer, minimum temperatures, and foliar sprays.

Model evaluation (Table 3) showed that predictions were reliable. For field-level yield, $R^2 = 0.63$, RMSE = 3.75, and d = 0.89; for flower-level yield, $R^2 = 0.70$, RMSE = 0.69, and d = 0.91, indicating robust performance. Observed versus predicted values are shown in Figure 2.

Saffron quality indices (Picrocrocin, Safranal, and Crocin content)

LASSO models were used to predict the content of picrocrocin, safranal, and crocin. Optimal λ values (λ_{min}) retained 7, 3, and 6 predictors for picrocrocin, safranal, and crocin, respectively (Table 4). For picrocrocin, the main negative predictors were altitude (-4.76) and number of foliar sprays (-4.41), while phosphorus fertilizer (3.97) positively influenced content. Minor contributions were observed from the minimum May temperature, planting depth, farm age, and minimum July temperature (Table 5). For safranal, altitude (quadratic, -4.82) and soil nitrogen (2.34) were the primary predictors, with planting depth showing minimal impact. For crocin, phosphorus fertilizer (18.3) had the strongest positive effect, whereas altitude (-12.99) was a moderate negative predictor. Smaller effects were observed for sulfur fertilizer, minimum July temperature, and number of irrigations.

Model performance (Table 6) showed moderate predictive ability. For picrocrocin, $R^2 = 0.24$ and RMSE = 8.71 (E1% at 257 nm); for safranal, $R^2 = 0.14$ and RMSE = 5.25 (E1% at 330 nm); and for crocin, $R^2 = 0.23$ and RMSE = 24.6 (E1% at 440 nm). The index of agreement ranged from 0.61 to 0.69, and mean errors were minor, indicating negligible bias. Observed versus predicted values are presented in Figure 3.

Overall, LASSO effectively identified key agronomic, climatic, and soil factors influencing saffron yield and quality. Phosphorus fertilization emerged as a potential influential factor for crocin and picrocrocin content, although the relatively low model performance suggests that this relationship should be interpreted with caution.

Table 1. Summary of cross-validation results for predicting stigma yield per hectare and stigma yield per kilogram of fresh flowers using LASSO regression models

	Stigma yield per hectare				***Stigma yield per kilogram of fresh flowers***			
Model type	**λ**	**MSE**	**SE**	**Number of Variables**	**λ**	**MSE**	**SE**	**Number of Variables**
λ_{1se}	0.8299	14.06	1.885	6	0.132	0.599	0.041	11

Note: λ: regularization parameters, MSE: mean squared error, SE: standard error.

Table 2. Estimated non-zero regression coefficients from the LASSO model for predicting stigma yield per hectare and stigma yield per kilogram of fresh flowers

Stigma yield per hectare				
Rank	**Predictor Variable**	**Transformation**	**Coefficient**	**Relative Importance**
-	Intercept	-	13.505	-
1	Amount of corms used ($kg.ha^{-1}$)	Linear	3.932	High
2	Manure fertilizer	Linear	1.838	Moderate
3	Average corm weight (g)	Quadratic (X^2)	0.697	Low
4	Field area	Linear	0.640	Low
5	Average corm weight (g)	Linear	0.539	Low
6	Irrigation frequency	Linear	0.172	Low
Stigma yield per kg of fresh flowers				
Rank	Predictor Variable	Transformation	Coefficient	Relative Importance
-	Intercept	-	12.469	-
1	Average corm weight (g)	Linear	0.429	High
2	Phosphorus fertilizer ($kg.ha^{-1}$)	Linear	0.319	High
3	Corm planting rate ($kg.ha^{-1}$)	Linear	0.300	High
4	Irrigation frequency	Quadratic (X^2)	0.216	Medium
5	Organic fertilizer ($ton.ha^{-1}$)	Linear	0.170	Medium
6	Max temperature in Dec (°C)	Quadratic (X^2)	0.138	Medium
7	Irrigation frequency	Linear	0.132	Medium
8	Nitrogen fertilizer ($kg.ha^{-1}$)	Quadratic (X^2)	-0.086	Low
9	Min temperature in Aug (°C)	Quadratic (X^2)	0.081	Low
10	Number of foliar sprays	Linear	0.080	Low
11	Min temperature in Mar (°C)	Quadratic (X^2)	0.034	Low

Note: "Quadratic (X^2)" indicates that the squared term of the variable was included in the model. Independent variables were standardized; Relative importance was qualitatively categorized based on absolute coefficient magnitude.

Discussion

The LASSO regression models identified a subset of key agronomic and environmental variables influencing saffron yield and quality, highlighting the dominant roles of phosphorus fertilization, corm characteristics, and irrigation management. Retaining a limited number of predictors supports model interpretability while maintaining acceptable predictive performance. Importantly, the results indicate that yield and quality responses are not uniformly driven by all measured variables, but rather by a smaller group of influential factors, while many others exhibited weak or negligible effects. This suggests that targeted management strategies focusing on high-impact variables may be more effective than uniform input application. Although regularization techniques improve model stability and reduce overfitting, the identified relationships should be interpreted with caution, particularly for predictors with lower standardized coefficients, as these may reflect context-specific or weak associations.

These findings are consistent with previous studies highlighting the effectiveness of regularization approaches in high-dimensional agricultural datasets (Tibshirani, 1996; Zou and Hastie, 2005; Friedman et al., 2010; Hastie et al., 2009). Although the number of predictors was relatively high compared to the number of observations, LASSO regression enabled effective variable selection and dimensionality reduction. Nevertheless, given the limited sample size, the results should be interpreted with caution, and further validation using larger datasets is recommended.

Table 3. Statistical indicators of LASSO model fit for saffron stigma yield per hectare and stigma yield per kg of fresh flower

Statistic	Define	Stigma yield per hectare	Stigma yield per kg of fresh flower
n	Number of observations	99	99
MeanObs	Mean observed value	13.51	12.47
MeanSim	Mean predicted value	13.51	12.47
RMSE	Root mean squared error	3.75	0.69
nRMSE (%)	Normalized RMSE	27.75	5.55
R^2	Coefficient of determination	0.63	0.70
d	Index of agreement	0.89	0.91
ME	Mean error	0.54	0.65
MSE	Mean squared error	14.04	0.48

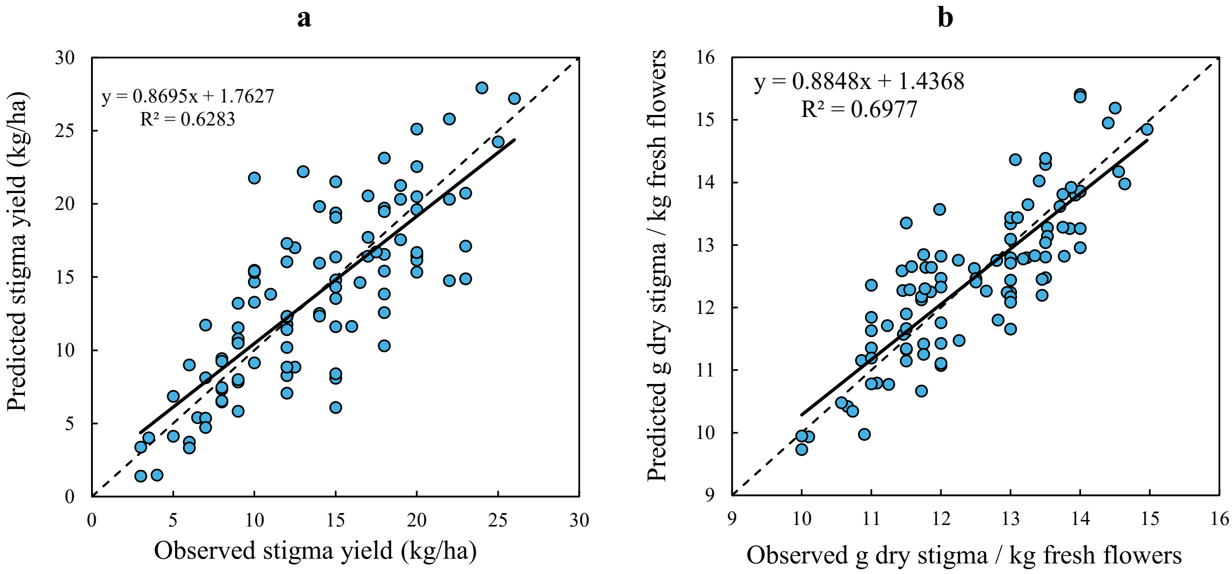

Figure 2. Observed versus predicted a) saffron stigma yield per hectare and b) stigma yield per kilogram of fresh flower.

Table 4. Summary of cross-validation results for predicting saffron quality indices (Picrocrocin, Safranal and Crocin content) using LASSO regression models

Model type	Picrocrocin content				Safranal content				Crocin content			
	λ	MSE	SE	Number of Variables	λ	MSE	SE	Number of Variables	λ	MSE	SE	Number of Variables
λ_min	1.04	61.7	10.5	7	0.665	11.9	3.19	3	2.64	273	25.4	6

Table 5. Estimated non-zero regression coefficients of the LASSO model at the optimal regularization parameter (λ_min) for predicting saffron quality indices (Picrocrocin, Safranal, and Crocin content).

Rank	Variable	Transformation	Coefficient	Relative Importance
Picrocrocin Content				
-	Intercept	-	80.9	-
1	Altitude	Linear	-4.76	High
2	Number of foliar sprays	Linear	-4.41	High
3	Phosphorus fertilizer ($kg.ha^{-1}$)	Linear	3.97	High
4	Min temperature in May (°C)	Quadratic (X^2)	3.21	Medium
5	Corm planting depth (cm)	Quadratic (X^2)	-1.47	Low
6	Farm age (Year)	Quadratic (X^2)	0.861	Low
7	Min temperature in Jul (°C)	Quadratic (X^2)	0.008	Low
Safranal Content				
-	Intercept	-	53.4	-
1	Altitude	Quadratic (X^2)	-4.82	High
2	Soil nitrogen	Quadratic (X^2)	2.34	Medium
3	Corm planting depth (cm)	Quadratic (X^2)	-0.085	Low
Crocin Content				
-	Intercept	-	53.4	-
1	Phosphorus fertilizer ($kg.ha^{-1}$)	Linear	18.3	High
2	Altitude	Linear	-12.99	Medium
3	Min temperature in Jul (°C)	Quadratic (X^2)	3.66	Low
4	Sulfur fertilizer ($kg.ha^{-1}$)	Linear	2.54	Low
5	Sulfur fertilizer ($kg.ha^{-1}$)	Quadratic (X^2)	1.96	Low
6	Number of irrigations	Quadratic (X^2)	0.265	Low

Note: "Quadratic (X^2)" indicates that the squared term of the variable was included in the model. Independent variables were standardized; Relative importance was qualitatively categorized based on absolute coefficient magnitude.

Table 6. Statistical indicators of LASSO model fit for saffron quality indices (Picrocrocin, Safranal, and Crocin content) in the optimal models

Statistic	Define	Picrocrocin	Safranal	Crocin
n	Number of observations	99	99	99
MeanObs	Mean observed value	80.9	53.43	183
MeanSim	Mean predicted value	80.9	53.43	183
RMSE	Root mean squared error	8.71	5.25	24.6
nRMSE (%)	Normalized RMSE	10.8	9.82	13.4
R^2	Coefficient of determination	0.24	0.14	0.23
d	Index of agreement	0.69	0.61	0.66
ME	Mean error	-0.26	-1.21	-1.06
MSE	Mean squared error	75.85	27.54	602

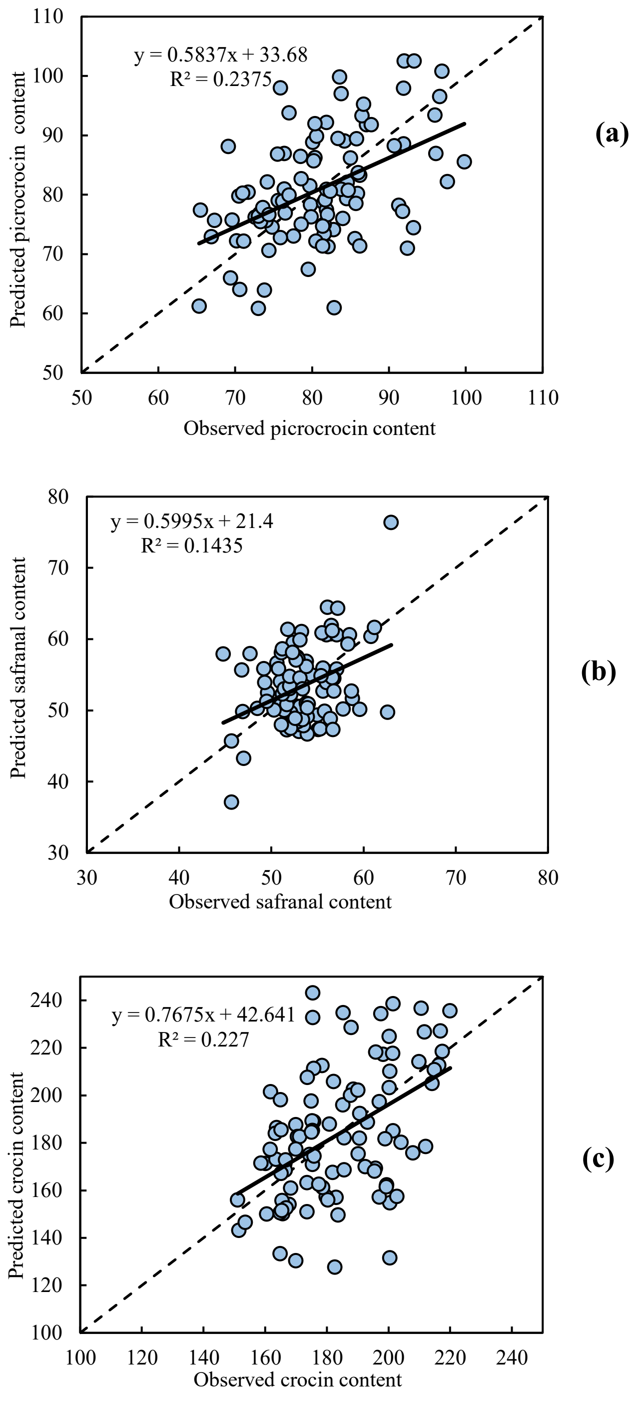

Figure 3. Observed versus predicted saffron picrocrocin (a), Safranal, and Crocin (c) content values for the optimal model.

For stigma yield per kilogram of fresh flowers, phosphorus fertilization, average corm weight, and planting density emerged as the most influential predictors, each showing a consistent positive effect. Organic manure and irrigation frequency had moderate positive contributions. The presence of a significant quadratic term for irrigation frequency suggests a nonlinear response, indicating that both insufficient and excessive irrigation may negatively affect flower quality. Excessive irrigation may reduce soil aeration and increase susceptibility to fungal diseases, thereby limiting crop performance (Kothari et al., 2021; Doostdar Mahmudabad and Azadi Brice, 2022).

Sulfur fertilizer exhibited weak and nonlinear effects, suggesting limited overall importance. The simultaneous inclusion of linear and quadratic terms for certain predictors reflects underlying curvilinear relationships, allowing the model to capture potential threshold or optimum responses while maintaining hierarchical consistency. These findings highlight the importance of optimizing input levels rather than maximizing them. The simultaneous inclusion of linear and quadratic terms for certain variables reflects nonlinear (curvilinear) relationships commonly observed in agronomic systems. Retaining both terms allows the model to capture potential threshold or optimum responses. This is also consistent with maintaining a hierarchical model structure.

At the field scale, stigma yield per hectare was primarily influenced by corm planting weight, followed by organic fertilization, corm weight, and field area, highlighting the importance of initial storage reserves, planting density, and soil fertility. The presence of a quadratic effect for corm weight indicates a nonlinear relationship with yield, suggesting that both undersized and oversized corms may limit productivity, likely due to insufficient reserves in smaller corms and intra-specific competition or physiological constraints in larger ones. The positive effects of organic fertilization and corm characteristics are consistent with previous studies that emphasize their roles in enhancing soil fertility and plant vigor (Rahimi et al., 2020; Askari & Keshavarz, 2021; Mardani et al., 2018). Similarly, increased planting density was associated with improved flower production, although this effect may depend on resource availability and should be interpreted within the context of local management conditions (Nassiri Mahallati et al., 2015).

Climatic factors also contributed to yield variation. Higher maximum temperatures during September and December were associated with reduced yield, potentially due to their influence on flower initiation and development. In contrast, moderate autumn temperatures and adequate early-season moisture appeared beneficial. However, these relationships may be context-dependent and should be interpreted with caution, given the use of interpolated climatic data.

Overall, the results suggest that saffron productivity is strongly influenced by a combination of planting material quality, nutrient management, and climatic conditions, with optimal input levels being more critical than maximum input application. Results were consistent with previous studies demonstrating that the performance of organic and mineral fertilizers was improved when applied appropriately (Koocheki and Seyyedi, 2019; Koocheki & Sabet Teimouri, 2014).

The analysis of saffron quality indices revealed trait-specific responses, with phosphorus fertilization consistently emerging as the strongest positive predictor of both picrocrocin and crocin content. In contrast, altitude showed a generally negative association with all three quality traits, while a significant quadratic effect for safranal suggests a mid-elevation optimum. Foliar spray frequency was negatively associated with picrocrocin, whereas soil nitrogen showed nonlinear effects on safranal, potentially reflecting variability in nitrogen use efficiency across conditions (Esmaealzadeh et al., 2023). Sulfur showed weak linear and quadratic effects, indicating a limited overall contribution. Temperature variables at critical growth stages also demonstrated nonlinear relationships with metabolite content (Khan et al., 2023), supporting the idea that saffron quality is sensitive to environmental variability. These patterns suggest that metabolite accumulation may be optimized under moderate nutrient availability and environmental conditions (Poudyal et al., 2020; Arora et al., 2024). However, this interpretation should be made with caution, as the relatively low R^2 values for quality traits indicate that the models explain only a limited portion of the observed variability. This implies that additional factors, such as genetic variation, post-harvest handling, and micro-environmental conditions, may play a substantial role but were not fully captured in the present analysis.

A key contribution of this study lies in prioritizing influential variables within a high-dimensional system and identifying nonlinear (threshold) responses, rather than merely detecting statistically significant factors. This approach provides a more nuanced representation of farm-level variability than conventional regression methods, though its applicability depends on data quality and model assumptions.

Overall, the results indicate that saffron yield and quality are governed by a complex interplay of agronomic practices, soil properties, and climatic conditions, with several variables exhibiting nonlinear or threshold effects. In particular, optimizing phosphorus fertilization, corm characteristics, and planting density appears critical for improving productivity, whereas suboptimal input management or extreme environmental conditions may reduce performance. These findings provide data-driven insights to support site-specific, climate-adaptive management strategies for saffron production under semi-arid conditions. However, given the relatively limited predictive performance for quality traits and the observational nature of the dataset, further validation with larger, more controlled datasets is required.

Conclusion

This study developed a data-driven framework to identify key agronomic, environmental, and management factors influencing saffron yield and quality in northeastern Iran using LASSO regression. From an initial set of 75 variables, only a few influential predictors were identified, highlighting the effectiveness of regularization for variable selection in high-dimensional agricultural systems. The results emphasize the importance of corm characteristics, planting density, and phosphorus management as primary drivers of yield and quality, while several variables exhibited nonlinear (threshold) responses, indicating that both insufficient and excessive inputs may reduce performance. Climatic conditions, particularly temperature and moisture during key growth stages, also contributed to yield and quality variability. These findings provide data-driven insights to support site-specific and climate-adaptive management strategies and can inform the development of decision-support systems (DSSs) for saffron production under semi-arid conditions. However, given the limited predictive performance for quality traits and the observational nature of the dataset, further validation with larger and more controlled datasets is recommended.

Acknowledgements

This research was partially supported by the Ferdowsi University of Mashhad (FUM) under Research Project No. 3.47925. The authors would like to express their sincere gratitude for this financial support.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

Arora, P.K., Tripathi, S., Omar, R.A., Chauhan, P., Sinhal, V.K., Singh, A., Srivastava, A., Garg, S.K., & Singh, V.P. (2024). Next-generation fertilizers: the impact of bionanofertilizers on sustainable agriculture. *Microb Cell Fact, 23*(1):254. https://doi.org/10.1186/s12934-024-02528-5.

Askari, Y., & Keshavarz, K. (2021). Study of planting density and corm size on saffron (*Crocus sativus* L.) yield in Yasuj climatic conditions. *Journal of Saffron and Medicinal Plants Extension, 3*(1), 51-44. (In Persian with English abstract).

Ayoubi, K. A., Eisvand, H. R., Heydari, S., & Mousavi-Fard, S. (2023). Effects of Foliar Application of Iron and Zinc Micronutrient Elements on Quantitative and Qualitative Yield of Saffron (*Crocus sativus* L.). *Journal of Saffron Research, 11*(1), 78-66. (In Persian with English abstract). https://doi.org/10.22077/jsr.2023.6083.1206.

Azizi, G., Moosavi, S. G., Seghatoleslami, M. J., & Fazeli Rostampour, M. (2023). Effect of different irrigation systems and fertilizer types on saffron corm production. *Journal of Medicinal Plants and By-Products, 12*(4), 339-348. https://doi.org/10.22092/jmpb.2022.356222.1410.

Bouyoucos, G. J. (1936). Directions for making mechanical analyses of soils by the hydrometer method. *Soil Science*, *42*(3), 225-230.

Bradstreet, R. B. (1954). Kjeldahl method for organic nitrogen. *Analytical Chemistry, 26*(1), 185-187.

Cardone, L., Castronuovo, D., Perniola, M., Scrano, L., Cicco, N., & Candido, V. (2020). The Influence of Soil Physical and Chemical Properties on Saffron (*Crocus sativus* L.) Growth, Yield, and Quality. *Agronomy, 10*(8), 1154. https://doi.org/10.3390/agronomy10081154.

Doostdar Mahmudabad, M., & Azadi Brice, S. (2022). Studying the effect of irrigation and drainage methods on root and crown rot disease of kiwi tree. *Seventh International Congress on Agricultural and Environmental Development with Emphasis on the United Nations Development Program.*

Esmaealzadeh D, Moodi Ghalibaf A, Shariati Rad M, Rezaee R, Razavi B. M., & Hosseinzadeh H. (2023). Pharmacological effects of Safranal: An updated review. *Iranian Journal of Basic Medical Sciences, 26*(10):1131-1143. https://doi.org/10.22038/IJBMS.2023.69824.15197.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01.

Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. Facta Universitatis, Series: Mathematics and Informatics, 583-595. https://doi.org/10.22190/FUMI1903583G.

Gresta F., Avola G., Lombardo G.M., Siracusa L., & Ruberto G. (2009). Analysis of flowering, stigmas yield and qualitative traits of saffron (*Crocus sativus* L.) as affected by environmental conditions. *Scientia Horticulturae, 119* (3): 320-324. https://doi.org/10.1016/j.scienta.2008.08.008

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer

Kaveh, H., & Salari, A. (2018). Study and comparison of saffron quality produced in major centers of production in Khorasan provinces. *Saffron Agronomy and Technology, 6*(2), 209-218. (In Persian with English abstract) https://doi: 10.22048/jsat.2017.59510.1184.

Khan, F., Siddique, A.B., Shabala, S., Zhou, M., & Zhao, C. (2023). Phosphorus plays key roles in regulating plants' physiological responses to abiotic stresses. *Plants, 12*(15): 2861. https://doi.org/10.3390/plants12152861.

Koocheki A., Seyyedi S.M. (2019). Mother corm origin and planting depth affect physiological responses in saffron (*Crocus sativus* L.) under controlled freezing conditions. *Industrial Crops and Products, 138*, 111468. https://doi.org/10.1016/j.indcrop.2019.111468.

Koocheki, A., & Sabet Teimouri, M. (2014). Effect of age of farm, corm size and manure fertilizer treatments on morphological criteria of Saffron (*Crocus sativus* L.) under Mashhad conditions. *Applied Field Crops Research, 27*(105). 148-157. https://doi.org/10.22092/aj.2014.103365.

Kothari, D., Thakur, M., Joshi, R., Kumar, A., & Kumar, R. (2021). Agro-climatic suitability evaluation for saffron production in areas of Western Himalaya. *Frontiers in Plant Science*, 12:657819. https://doi.org/10.3389/fpls.2021.657819.

Lozano P., Delgado D., Gómez D., Rubio M., & Iborra J.L. (2000). A non-destructive method to determine the safranal content of saffron (*Crocus sativus* L.) by supercritical carbon dioxide extraction combined with high-performance liquid chromatography and gas chromatography. *Journal of Biochem Biophys Methods, 43*(1-3):367-78. https://doi.org/10.1016/s0165-022x(00)00090-7.

Mardani Asl, S.A., Movahhedi Dehnavi, M., Salehi, A., & Yadavi, A. (2018). Effect of corm weight and planting density on saffron (*Crocus sativus* L.) yield between apple trees in Firouzabad of Yasouj. *Journal of Saffron Research, 6*(1), 89-102. (In Persian with English abstract) https://doi.org/10.22077/jsr.2017.823.1034.

Mohammadkhani, F., Pouryousef, M., & Yousefi, A.R. (2023). Growth and production response in saffron-chickpea intercropping under different irrigation regimes. *Industrial Crops and Products, 193*: 116256. https://doi.org/10.1016/j.indcrop.2023.116256.

Muche M., Kokeb A., & Molla E. (2015). Assessing the physicochemical properties of soil under different land use types. *Journal of Environmental & Analytical Toxicology, 5*, 309. https://doi.org/10.4172/2161-0525.1000309.

Nassiri Mahallati, M., Koocheki, A., Amin Ghafouri, A. & Mahluji Rad, M. (2015). Optimizing corm size and density in saffron (*Crocus sativus* L.) cultivation by central composite design. *Saffron Agronomy and Technology, 3*(3), 161-177. (In Persian with English abstract) https://doi.org/10.22048/jsat.2015.10386.

Nelson, D.W., & Sommers, L. E. (1996). Total carbon, organic carbon, and organic matter. In *Methods of Soil Analysis; Part 3*, SSSA Book Series; Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R. H., Soltanpour, P.N., Tabatabai, M. A., Johnston, C. T., Sumner, M. E., Eds.; Madison: Indianapolis, IN, USA, 1996; pp. 961–1010.

Olsen, S. R. (1954). *Estimation of Available Phosphorus in Soils by Extraction with Sodium Bicarbonate (No. 939).* US Department of Agriculture.

Pirasteh-Anosheh, H., Babaie-Zarch, M. J., Nasrabadi, M., Parnian, A., Alavi-Siney, S. M., Beyrami, H., Kaveh, H., Hashemi, S. E., Durrer, U., McDonald, K., & Race, M. (2023). Climate and management factors influence saffron yield in different environments. *Agrosystems, Geosciences & Environment, 6*, e20418. https://doi.org/10.1002/agg2.20418 .

Poudyal, S., Owen, J. S., Sharkey, T. D., Fernandez, R. T., & Cregg, B. (2020). Phosphorus requirement for biomass accumulation is higher compared to photosynthetic biochemistry for three ornamental shrubs. *Scientia Horticulturae, 275,* 109719. https://doi.org/10.1016/j.scienta.2020.109719.

Riyahi, R., Mirdamadi, S., Farrajollah Hosseini, S., & Omidi Najafabadi, M. (2023). Investigating the role of saffron business clusters in rural development (Case Study: South Khorasan Province). *Village and Development, 25*(4), 213-231. https://doi.org/10.30490/rvt.2023.357120.1432.

Sepaskhah, A., & Kamgar-Haghighi, A. (2012). Saffron Irrigation Regime. *International Journal of Plant Production, 3*(1), 1-16. https://doi.org/10.22069/ijpp.2012.627.

Shahnoushi N., Abolhassani L., Kavakebi V., Reed M., & Saghaian S. (2020). Economic analysis of saffron production. *Saffron: Science, Technology and Health*. pp. 337-356. https://doi.org/10.1016/B978-0-12-818638-1.00021-6

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B, 58*(1), 267–288.

Walkley, A. J., & Black, I. A. (1934) Estimation of soil organic carbon by the chromic acid titration method. *Soil Science, 37*, 29–38.

Wang Z., Li X., Xu J., Yang Z., & Zhang Y. (2021). Effects of ambient temperature on flower initiation and flowering in saffron (*Crocus sativus* L.). *Scientia Horticulturae. 279*, 109859. https://doi.org/10.1016/j.scienta.2020.109859.

Wang, X.J., Jia, Zh. K., Liang, L.Y., & Kang, Sh. Zh. (2013). Effect of manure management on the temporal variations of dryland soil moisture and water use efficiency of Maize. *Journal of Agricultural Science and Technology, 15*, 1293-1304.

Willmott, C.J. (1981). On the validation of models. *Physical Geography, 2*(2), 184–194.

Ziaei, S.M., Feizi, H., Khashei Siuki, A. & Sahabi, H. (2024). Yield and quality of saffron (*Crocus sativus* L.) in response to priming treatments and water deficit. *Italian Journal of Agronomy. 19*(3): 100020. https://doi.org/10.1016/j.ijagro.2024.100020.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 67*(2): 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.