Predicting AquaCrop-Simulated Durum Wheat Yield with Machine Learning: Algorithm Comparison and Agronomic Signal Convergence in the Capitanata Plain

Pasquale Garofalo *, Anna Rita Bernadette Cammerino, Maria Riccardi

Posted Date: 20 March 2026

doi: 10.20944/preprints202603.1628.v1

Keywords: *Triticum durum*; AquaCrop; multilayer perceptron; support vector regression; decision tree; yield prediction; Mediterranean agriculture

 doi:10.20944/preprints202603.1628.v1

Article

Predicting AquaCrop-Simulated Durum Wheat Yield with Machine Learning: Algorithm Comparison and Agronomic Signal Convergence in the Capitanata Plain

Pasquale Garofalo [1,*], Anna Rita Bernadette Cammerino [2] and Maria Riccardi [3]

[1] Council for Agricultural Research and Economics - Agriculture and Environment (CREA-AA), Via Celso Ulpiani 5, 70125 Bari, Italy

[2] Department of Science of Agriculture, Food, Natural Resources and Engineering, University of Foggia, Via Napoli 25, 71122 Foggia, Italy

[3] National Research Council of Italy - Institute for Agricultural and Forestry Systems in the Mediterranean (CNR-ISAFoM), 80055 Portici, NA, Italy

* Correspondence: pasquale.garofalo@crea.gov.it

Abstract

Five machine learning algorithms — Linear Regression (LR), Multilayer Perceptron (MLP), Support Vector Machine for regression (SMOreg), RandomTree, and Reduced Error Pruning Tree (REPTree) — were trained and compared for predicting durum wheat (*Triticum durum* Desf.) grain yield simulated by AquaCrop-GIS across the Capitanata plain (Southern Italy). A dataset of 342 instances was constructed by crossing 25 soil profiles, three sowing dates, and two irrigation regimes over 15 climatic grid cells (2014–2023), validated by stratified 10-fold cross-validation. MLP achieved the highest accuracy ($R = 0.983$; MAE = 0.059 t ha^{-1}; RMSE = 0.083 t ha^{-1}); the four interpretable models clustered at $R = 0.891$–0.907 (RMSE = 0.192–0.203 t ha^{-1}). All models converged on consistent agronomic signals: standard sowing (1 November) yielded +0.53 t ha^{-1} over late sowing (15 November); supplemental irrigation added +0.17 t ha^{-1}; high-silt and clay soils produced superior yields. The SMOreg normalised weight vector identified autumn temperature (Tmin_oct_nov: −0.462; Tmax_oct_nov: −0.405) as the dominant climate predictor, reflecting the AquaCrop phenological mechanism whereby elevated early-season thermal loads curtail tillering. The convergence of directional signals across fundamentally different algorithmic architectures — linear, kernel-based, and tree-based — confirms that ML surrogates can efficiently emulate AquaCrop response surfaces for scenario analysis and decision-support in Mediterranean dryland farming systems.

Keywords: *Triticum durum*; AquaCrop; multilayer perceptron; support vector regression; decision tree; yield prediction; Mediterranean agriculture

1. Introduction

Durum wheat (*Triticum durum* Desf.) represents one of the most economically and ecologically significant cereal crops in the Mediterranean basin, providing the principal raw material for pasta, couscous, and semolina production worldwide [1]. Within the European Union, Italy is the leading producer, accounting for approximately 38% of the EU durum wheat area, with the Capitanata plain (Foggia Province, Apulia) standing as one of the most productive zones on the continent [2]. The region is characterised by heterogeneous soils ranging from sandy-loam Inceptisols in the eastern coastal sectors to deep clay-rich Vertisols in the central plain, and by management strategies spanning from fully rainfed continuous monocultures to systems with targeted supplemental irrigation at anthesis.

 doi:10.20944/preprints202603.1628.v1

Climate change is progressively altering the thermal and hydrological boundary conditions of Mediterranean cereal production: rising spring temperatures, increased frequency of late-season drought episodes, and greater interannual precipitation variability are projected to reduce yield potential and amplify production risk in the coming decades [3]. In response, process-based crop models have become indispensable research tools for evaluating the sensitivity of cropping systems to soil-climate-management combinations at a spatial resolution and temporal depth that field experimentation alone cannot achieve. AquaCrop [4,5], developed by the Food and Agriculture Organization of the United Nations (FAO), is one of the most widely adopted such models for water-limited environments. It simulates canopy development, biomass accumulation, and grain yield as functions of thermal time, soil water balance, and plant physiological parameters, and has been validated for a wide range of crops in water-limited Mediterranean environments [6,7]. At regional scale, Aqua-Crop-GIS has been applied by Garofalo and Cammerino [8] to map yield productivity, carbon footprint, and water footprint of durum wheat across the Capitanata plain under different sowing dates and irrigation regimes, providing the foundational simulation dataset for the present study.

Machine learning (ML) algorithms have emerged over the past decade as a powerful complement to process-based crop models. By training on model outputs -a paradigm known as surrogate or emulator modelling -ML can reproduce complex multi-variate input-output relationships at negligible computational cost, enabling rapid scenario exploration across parameter spaces that would be prohibitively expensive to simulate exhaustively [9,10]. Studies covering a wide range of environments and crops consistently identify temperature, soil characteristics, and management variables as the most influential predictors, and report that non-linear models such as Random Forest, SVM, and ANN outperform linear regression baselines in the majority of applications [9–13]. Shahhosseini et al. [14] achieved R^2 up to 0.725 when coupling ML with the APSIM crop model for maize in the US Corn Belt, with soil water-related simulation variables identified as the most influential inputs. Gautam et al. [15] compared AquaCrop, a semi-physical model, and ANN for corn and soybean yield estimation in Illinois over 25 years, finding that the ANN achieved the highest accuracy (R^2 = 0.96 for soybean; R^2 = 0.89 for corn), while AquaCrop provided mechanistically interpretable predictions. For wheat specifically, Bassu et al. [16] demonstrated that optimised sowing date in a variable Mediterranean environment can substantially reduce yield risk, a management principle directly relevant to the present study.

Within the ML family, the comparison between interpretable and black-box models raises a fundamental practical question for agronomic applications. Multilayer Perceptrons (MLP) and other neural architectures [17] can approximate arbitrarily complex non-linear functions and typically achieve the highest predictive accuracy, but their distributed weight structure prevents direct attribution of predictions to individual input variables, limiting their utility for deriving agronomic insights or communicating results to practitioners and policy-makers. Decision trees -in particular pruned variants such as Reduced Error Pruning Tree (REPTree) -offer transparent, rule-based structures that agronomists and decision-makers can directly interpret, at a moderate cost in predictive accuracy. Fernandez-Delgado et al. [18] evaluated 179 classifiers across 121 benchmark datasets and showed that random forest-based algorithms tend to outperform other methods across diverse problem types. Support Vector Machines for regression (SMOreg) occupy an intermediate interpretability-accuracy position: with a linear kernel, the normalised weight vector provides a direct comparative importance ranking across predictors on a common dimensionless scale [19], making it a useful tool for agronomic signal extraction.

The recent integration of AI and remote sensing in Mediterranean agroecosystems [20] highlights the growing demand for algorithm-level comparisons that go beyond accuracy metrics to examine the agronomic interpretability of ML outputs. Despite this growing interest, systematic comparisons spanning linear, kernel-based, and tree-based algorithms on the same AquaCrop-derived dataset for durum wheat in a Mediterranean sub-regional context remain limited. Most existing comparative studies either focus on field-scale observational data -where measurement noise

 doi:10.20944/preprints202603.1628.v1

and spatial variability cap achievable accuracy -or employ a restricted subset of algorithms without the full spectrum from linear to non-linear interpretable to black-box. Furthermore, the agronomic interpretation of ML model outputs is rarely the central focus of yield prediction papers: predictor importance is often reported as a secondary result without systematic cross-model convergence analysis. The present study addresses these gaps by comparing five ML algorithms -Linear Regression (LR), Multilayer Perceptron (MLP), SMOreg, RandomTree, and REPTree -trained on a dataset of 342 instances derived from AquaCrop-GIS simulations across the Capitanata plain. The specific objectives were: (i) to evaluate and compare the predictive accuracy of the five algorithms using 10-fold cross-validation; (ii) to identify the most agronomically relevant predictors through model-specific importance measures including regression coefficients, normalised SVM weights, and tree split importance; and (iii) to assess the convergence of agronomic signals across algorithms as a form of algorithm-independent validation of management recommendations.

2. Materials and Methods

2.1. Study Area

The study area is the Capitanata plain, Foggia Province, Apulia, Southern Italy (41.0°-41.8°N, 15.2°-16.1°E; 50-200 m a.s.l.), Italy's leading durum wheat district accounting for approximately 38% of total national cultivated area [2,8]. The climate is sub-humid Mediterranean (Köppen BSk-Csa transition), with mean annual precipitation of 450-600 mm concentrated between October and April, and mean annual temperature of approximately 15 °C. The central plain receives around 460-530 mm yr^{-1} while the western hilly fringe receives up to 600-630 mm yr^{-1} [8]. Soils are predominantly Vertisols and Inceptisols, with clay content ranging from 18% to 45%, bulk densities of 1.30-1.50 g cm^{-3}, and Plant Available Water (PAW) between 115 and 210 mm; sandy-loam Entisols with PAW < 50 mm occur in the eastern coastal sectors [8]. Durum wheat is typically sown between mid-October and mid-November and harvested in June-July, with anthesis occurring in late April-early May. The area is characterised by increasing water and thermal stress associated with ongoing climate change [3].

2.2. AquaCrop Simulations and Dataset Construction

Yield simulations were performed using AquaCrop-GIS (GIS extension) [4,5], parameterised for durum wheat following Garofalo and Cammerino [8], who calibrated crop coefficients against the Capitanata growing conditions. Simulations covered the period 2014-2023 (10 cropping seasons), crossing three sowing dates -early sowing (*ant*: 15 October), standard sowing (*reg*: 1 November), and late sowing (*post*: 15 November) -with two irrigation treatments: rainfed (No) and supplemental irrigation at anthesis (Flow, ~50 mm). Daily climatic inputs (maximum and minimum temperature, precipitation, reference evapotranspiration ET_0) for 15 grid cells (25 × 25 km) were sourced from the Agri4Cast/JRC database [22]. Reference evapotranspiration was calculated with the FAO Penman-Monteith equation [23]. Twenty-five soil profiles from the Harmonised World Soil Database v2.0 [24] were intersected with the climatic grid, yielding 57 unique pedo-climatic combinations.

For each combination of pedo-climatic cell × sowing date × irrigation treatment, AquaCrop was run across 10 simulation years and the output grain yield (t ha^{-1}, at 0% moisture) was averaged to produce a single representative productivity value integrating local climate variability. This averaging approach, consistent with Garofalo and Cammerino [8], reduces the effect of exceptional years and provides an estimate of the long-run yield potential for each soil-climate-management combination. The final dataset comprises 342 instances (from 57 pedo-climatic cell combinations × 3 sowing dates × 2 irrigation treatments).

2.3. Description of Predictor Variables

The 22 predictors are organised into three categories. Soil variables (4): sand (S_SAND, %), clay (S_CLAY, %), silt (S_SILT, %), and bulk density (S_BULK_DEN, g cm^{-3}). Management variables (2):

 doi:10.20944/preprints202603.1628.v1

sowing date (sow_time: ant / reg / post, nominal) and irrigation regime (Irr: No / Flow, nominal). Climate variables (16): cumulated values over four bimonthly windows (October-November, December-January, February-March, April-May), covering maximum temperature (Tmax, °C·10d^{-1}), minimum temperature (Tmin, °C·10d^{-1}), precipitation (Rain, mm), and reference evapotranspiration (ET_0, mm). Cumulated rather than mean seasonal values were used to preserve the temporal dynamics of crop water and heat demand across phenological stages.

2.4. Machine Learning Algorithms

All five algorithms were implemented and evaluated under a common framework. Nominal predictors (sow_time, Irr) were internally binarised using one-hot encoding prior to model training; numerical predictors were normalised to [0,1] for SMOreg. Predictive performance was estimated by stratified 10-fold cross-validation [25]. The algorithm implementations and hyperparameter settings adopted are described below.

2.4.1. Linear Regression (LR)

Linear Regression fits a hyperplane through the predictor space by minimising the residual sum of squares, subject to a ridge (L_2) regularisation penalty. The model takes the general form:

$$\hat{y} = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \varepsilon \quad (1)$$

where $\hat{y}$ is the predicted yield (t ha^{-1}), β_0 is the intercept, β_i are the fitted regression coefficients for predictors x_i, and ε is the residual error. Predictor selection was performed with an AIC-based stepwise procedure: the Akaike Information Criterion penalises model complexity proportionally to the number of retained predictors, so a predictor is included only when the gain in fit exceeds its complexity cost. The ridge (L_2) regularisation parameter λ was set to 10^{-8}, a near-zero value that reproduces ordinary least-squares behaviour in practice while preventing numerical instability from multicollinearity among correlated seasonal climate predictors. The fitted coefficients, organised by predictor category, are illustrated in Figure 1.

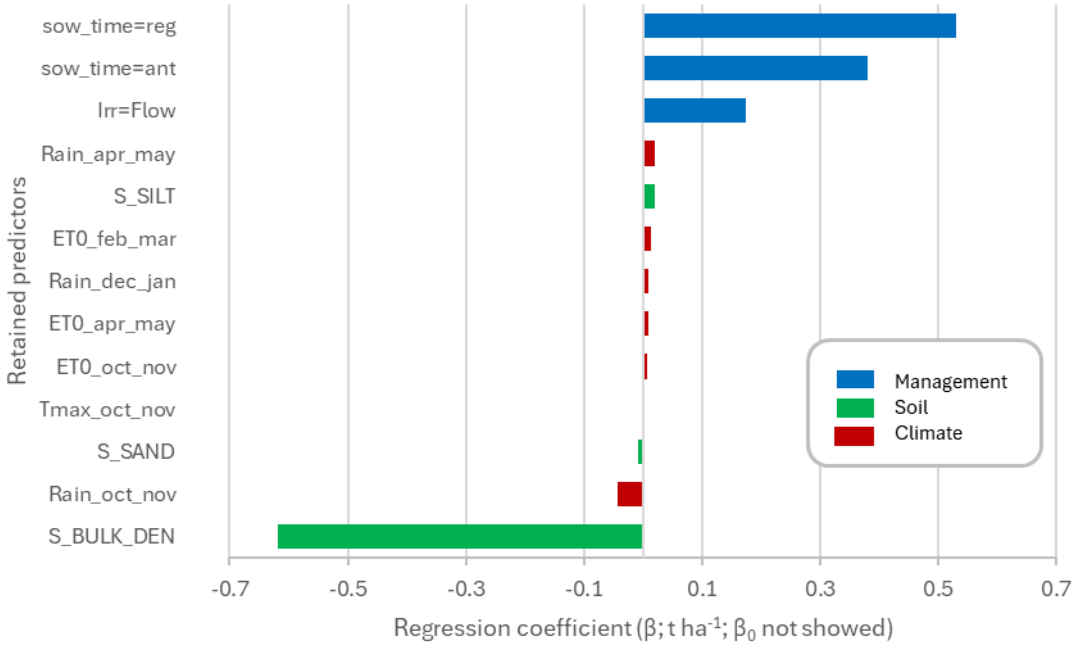

Figure 1. Fitted Linear Regression coefficients (β) for the 13 predictors retained by AIC-based stepwise selection (intercept β_0 = 5.305 t ha^{-1}). Bars are colour-coded by predictor category: blue = management, green = soil, red = climate. Positive values denote a yield-increasing effect; negative values a yield-decreasing effect.

 doi:10.20944/preprints202603.1628.v1

Criterion (AIC)-based stepwise procedure, and the ridge parameter λ was set to 10^{-8}. The LR model provides the most direct agronomic interpretation, since each retained β_i quantifies the marginal contribution of predictor x_i in original units (t ha^{-1} per unit of predictor), holding all other variables constant. The ridge penalty is equivalent to minimising:

$$\sum_i (y_i - \hat{y}_i)^2 + \lambda \sum_i \beta_i^2 \quad (2)$$

which shrinks large coefficients toward zero, reducing the risk of multicollinearity inflation in correlated climate predictors.

2.4.2. Multilayer Perceptron (MLP)

The MLP [17] is a feedforward artificial neural network with one hidden layer comprising 12 nodes, computed as H = [(p + 1)/2] where p = 22 is the number of input features and 1 is the output, yielding H = 12 hidden units. The network transforms inputs through successive non-linear projections:

$$h = \sigma(W^1 x + b^1) \quad (3)$$

$$\hat{y} = W^2 h + b^2 \quad (4)$$

where x is the input vector, h is the hidden-layer activation vector, W^1, W^2 are weight matrices, b^1, b^2 are bias vectors, and σ is the sigmoid activation function applied element-wise:

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (5)$$

Training used back-propagation with learning rate $\eta = 0.3$, momentum $\alpha = 0.2$, and 500 epochs. The learning rate determines the magnitude of each weight update: a value of 0.3 is a well-validated default that ensures steady convergence without oscillation on datasets of this size. Momentum (α = 0.2) carries a fraction of the previous weight increment into the current step, accelerating progress along shallow gradient directions and helping escape flat loss regions; the conservative value was chosen to avoid destabilising training on n = 342 instances. Five hundred epochs were sufficient for convergence given the modest number of trainable weights (313) and the smooth, deterministic nature of the AquaCrop response surface. The sigmoid activation ensures that each hidden node can capture non-linear transformations of the input space, allowing the MLP to approximate any continuous function to arbitrary precision (universal approximation theorem). However, the distributed nature of the weight matrices W^1 and W^2 means that individual predictor importance cannot be directly extracted: post-hoc methods such as SHAP values [26] would be required for local attribution, which is beyond the scope of the present study.

2.4.3. Support Vector Machine for Regression (SMOreg)

SMOreg implements the ε-insensitive Support Vector Regression (SVR) optimisation [19] using the Sequential Minimal Optimisation (SMO) algorithm. With a linear kernel, SVR finds the weight vector w and bias b that minimise:

$$\min \frac{1}{2} \| w \|^2 + C \sum_i (\xi_i + \xi_i^*) \quad (6)$$

subject to: $y_i - (w^T x_i + b) \leq \varepsilon + \xi_i$ and $(w^T x_i + b) - y_i \leq \varepsilon + \xi_i^*$, with $\xi_i, \xi_i^* \geq 0$, where C = 1.0 is the regularisation constant, ε = 0.001 is the tube half-width, and ξ_i, ξ_i^* are slack variables. The linear kernel $K(x_i, x_i') = x_i^T x_i'$ was adopted to produce a single normalised weight w_j for each predictor j, directly comparable across variables and interpretable as a relative importance ranking. The regularisation constant C = 1.0 governs the trade-off between margin width and training error: a unit value enforces balanced penalisation that avoids both under-fitting and excessive sensitivity to individual instances. The tube half-width ε = 0.001 defines the insensitive zone within which prediction errors incur no penalty: because training data are generated by a deterministic simulator, a near-zero ε is appropriate and allows the model to fit the response surface closely without unnecessary slack-variable costs. The linear kernel was preferred over radial basis or polynomial alternatives because it yields a single global weight vector — directly interpretable as a continuous importance ranking across all predictors on a common dimensionless scale — at no loss of accuracy relative to non-linear kernels on this structured dataset. The margin-based Lε loss function makes SMOreg more robust to outliers

 doi:10.20944/preprints202603.1628.v1

than ordinary least squares, and its support-vector mechanism avoids overfitting on the moderate-size dataset (n = 342).

2.4.4. RandomTree

RandomTree [27] builds an unpruned regression tree using the variance-reduction splitting criterion of Classification and Regression Trees, with the addition of random attribute selection at each node. At each node t, a random subset of K = √p candidate features is evaluated; the feature x_j and threshold θ maximising variance reduction are selected:

$$\Delta I(t) = \mathrm{Var}(y|t) - \frac{|t_L|}{|t|}\mathrm{Var}(y|t_L) - \frac{|t_R|}{|t|}\mathrm{Var}(y|t_R) \quad (7)$$

where t_L, t_R are the left and right child nodes and Var(·) is the response variance. Algorithm settings: maximum depth 10, minimum leaf instances m = 1.0, minimum variance proportion δ = 0.001. The resulting tree has 398 nodes, fitting the AquaCrop response surface at fine granularity. The random feature selection introduces implicit regularisation that reduces correlation between successive splits, at the cost of some variance compared to a greedy full-feature tree.

2.4.5. Reduced Error Pruning Tree (REPTree)

REPTree builds a full variance-reduction tree and then applies reduced-error post-pruning [25]: branches whose removal does not increase the mean squared error on a 33% held-out pruning subset are recursively eliminated. The pruning criterion replaces each pruned subtree with the leaf predicting the mean response of the corresponding training instances. The resulting tree has 73 nodes -approximately 18% of the RandomTree size -while achieving a RMSE only 0.004 t ha^{-1} higher, demonstrating the substantial compressibility of the AquaCrop response surface once the dominant variance structure is captured by the first few splits. Algorithm settings: 3-fold internal CV for pruning, minimum leaf instances m = 2, minimum variance fraction δ = 0.001. The 3-fold internal cross-validation for pruning reserves one third of each training fold as a dedicated pruning validation set, so the branch-removal decisions are driven by held-out prediction error rather than a fixed depth limit. The minimum leaf size m = 2 (one instance more than RandomTree's m = 1) imposes a mild smoothing constraint consistent with the pruning objective, preventing singleton leaves that might survive pruning by chance while preserving well-supported partitions. The minimum variance fraction δ = 0.001 suppresses any split that explains less than 0.1% of the residual response variance, eliminating negligible branches before the pruning pass is applied.

2.5. Model Evaluation and Validation Protocol

All algorithms were evaluated with 10-fold cross-validation (CV): the 342 instances were randomly partitioned into 10 equal folds, with 9 folds used for training and the remaining fold for testing in rotation. This procedure provides an unbiased estimate of generalisation performance and is the standard protocol for ML algorithm comparison [25]. Performance was assessed using five metrics:

Pearson correlation coefficient R (dimensionless):

$$R = \frac{\sum(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{\sum(y_i-\bar{y})^2 \cdot \sum(\hat{y}_i-\bar{\hat{y}})^2}} \quad (8)$$

Mean Absolute Error (MAE, t ha^{-1}):

$$\mathrm{MAE} = \frac{1}{n}\sum_i |y_i - \hat{y}_i| \quad (9)$$

Root Mean Squared Error (RMSE, t ha^{-1}):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2} \quad (10)$$

Relative Absolute Error (RAE, %) and Root Relative Squared Error (RRSE, %) normalise the absolute errors by the performance of a baseline mean predictor, facilitating comparison across studies with different yield ranges.

 doi:10.20944/preprints202603.1628.v1

3. Results

3.1. Overall Cross-Validation Performance

Table 1 summarises the 10-fold cross-validation metrics for all five algorithms. Predicted yield ranged from 3.95 t ha^{-1} (late sowing, rainfed, low-ET$_0$, low-silt) to 6.42 t ha^{-1} (early sowing, S_SILT ≥ 43.5%), a range of 2.47 t ha^{-1}, consistent with the variation reported by Garofalo and Cammerino [8] for the same study area.

The MLP achieved the highest accuracy on all five metrics (R = 0.983; MAE = 0.059 t ha^{-1}; RMSE = 0.083 t ha^{-1}; RAE = 16.82%; RRSE = 18.47%), with a RMSE equivalent to approximately 3.4% of the observed yield range. The remaining four algorithms clustered in a narrow band: R ranged from 0.891 (LR) to 0.907 (SMOreg) and RMSE from 0.192 (SMOreg) to 0.203 t ha^{-1} (LR), a range of only 0.011 t ha^{-1}. Despite their lower R relative to MLP, all interpretable models produced errors that are modest relative to the observed yield range, confirming that each model captures the dominant sources of variance in the dataset.

Table 1. Cross-validation performance metrics for the five ML algorithms (n = 342; 10-fold CV). Best value per column in bold.

Algorithm	R	MAE (t ha^{-1})	RMSE (t ha^{-1})	RAE (%)	RRSE (%)
Multilayer Perceptron (MLP)	**0.983**	**0.059**	**0.083**	**16.82**	**18.47**
SMOreg	0.907	0.136	0.192	38.70	42.81
RandomTree	0.903	0.112	0.194	32.03	43.24
REPTree	0.898	0.139	0.198	39.81	44.16
Linear Regression (LR)	0.891	0.151	0.203	43.21	45.24

3.2. Linear Regression: Coefficient Analysis

Of the five algorithms compared, LR was at once the simplest and the least accurate, yet it still captured approximately 79% of the variance in simulated yield (R = 0.891, RMSE = 0.203 t ha^{-1}, MAE = 0.151 t ha^{-1}). Its coefficient structure offered the most immediate agronomic readability of all models tested. The AIC-based stepwise selector retained 13 of 22 predictors (Figure 1); the nine excluded predictors (S_CLAY, Rain_feb_mar, Tmin_oct_nov, Tmax_dec_jan, Tmin_dec_jan, Tmin_feb_mar, Tmax_feb_mar, Tmin_apr_may, ET$_0$_dec_jan) did not reach significance under the regularisation penalty, indicating collinearity or marginal incremental contribution once the dominant predictors were fitted.

Management variables exerted the largest absolute effects. The sowing-date predictor was entered as two binary dummies relative to the base category (post, 15 November): a shared coefficient β = +0.380 t ha^{-1} applies to all earlier sowing dates (ant and reg), and an additional β = +0.151 t ha^{-1} is added exclusively for standard sowing (reg). The net yield advantages relative to late sowing are therefore: reg vs. post: +0.531 t ha^{-1}; ant vs. post: +0.380 t ha^{-1}. Supplemental irrigation (Irr = Flow) added +0.174 t ha^{-1}. Among soil variables, bulk density had the strongest negative coefficient (−0.618 t ha^{-1} per g cm^{-3}), reflecting the compaction-induced reduction of plant-available water in the AquaCrop framework. Silt contributed positively (+0.021 t ha^{-1} per %), while sand had a small negative effect (−0.008 t ha^{-1} per %). Spring rainfall (Rain_apr_may) carried a positive coefficient (+0.021 t ha^{-1} per mm), while autumn rainfall (Rain_oct_nov) was negatively associated (−0.042 t ha^{-1} per mm), possibly reflecting waterlogging risk during crop establishment in heavy soils. ET$_0$ variables in spring and late winter were positively associated with yield, consistent with the role of potential evapotranspiration as a proxy for radiation and vapour pressure deficit.

3.3. Multilayer Perceptron

 doi:10.20944/preprints202603.1628.v1

The MLP stood clearly apart from all other algorithms, compressing prediction errors to a RMSE of just 0.083 t ha^{-1} — roughly 2.4 times lower than the next best model and equivalent to only 3.4% of the observed yield range (R = 0.983, MAE = 0.059 t ha^{-1}). The 12-node sigmoid network (500 epochs) approximated the deterministic AquaCrop response surface with high fidelity, consistent with findings reported by Shahhosseini et al. [14] and Gautam et al. [15] for comparable ML-crop model coupling studies. The network architecture comprised 24 input units (after one-hot encoding of nominal predictors), 12 hidden units, and 1 output unit, totalling 24×12 + 12 + 12×1 + 1 = 313 trainable parameters for 342 training instances -a relatively constrained parameterisation that limits overfitting. Being a distributed representation, the MLP weight matrices (Equations 3-4) do not permit direct attribution of individual predictions to specific predictors. Their value in the present study lies primarily in establishing an upper bound on achievable accuracy for the given dataset, against which the interpretable models are benchmarked. The scatter plot in Figure 2 confirms this: data points for the MLP fall tightly along the 1:1 reference line across the full yield range (3.95–6.42 t ha^{-1}), with residuals distributed symmetrically around zero and no systematic bias at the extremes. The LR predictions shown in the same panel display the broader, more structured scatter characteristic of a linear model applied to a partially non-linear response surface, with a tendency to over-predict at low yields and marginally under-predict at the highest yields.

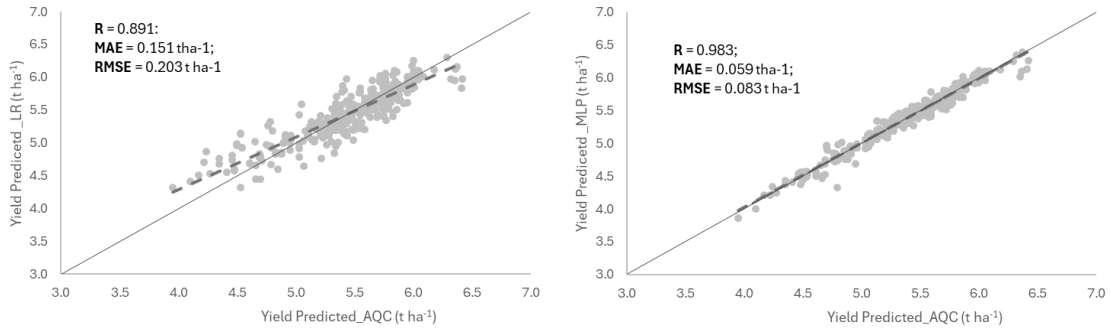

Figure 2. Observed versus predicted durum wheat grain yield (t ha^{-1}) for Linear Regression (LR) and Multilayer Perceptron (MLP) from 10-fold cross-validation (n = 342). The dashed line is the 1:1 reference; solid lines are model-specific regression fits.

3.4. SMOreg: Normalised Weight Analysis

Ranking second among the interpretable models (R = 0.907, RMSE = 0.192 t ha^{-1}, MAE = 0.136 t ha^{-1}), SMOreg offered a particularly useful analytical vantage point: unlike tree-based models, whose structure reflects variance-partitioning thresholds, the linear-kernel weight vector provides a continuous importance signal across all 22 predictors on a common dimensionless scale (Fig. 3). The four largest absolute weights are associated with temperature variables: Tmin_oct_nov (−0.462) and Tmax_oct_nov (−0.405) dominate as the two most important predictors with negative effects, while Tmin_feb_mar (+0.385) and Tmax_apr_may (+0.366) carry the largest positive weights. The strong negative autumn temperature signal reflects the AquaCrop phenological mechanism: higher cumulated thermal loads during October-November accelerate vegetative development and curtail the tillering phase, reducing the number of productive tillers and ultimately spike number per unit area. The positive effect of late-winter minimum temperature (Tmin_feb_mar) reflects reduced frost damage risk at the booting stage, while higher spring maximum temperature (Tmax_apr_may) in the Capitanata context is associated with radiation-driven grain-filling conditions prior to the onset of summer heat extremes.

Among soil predictors, S_SILT (+0.230) had the largest positive weight, while S_SAND (−0.143) had the largest negative soil weight. Management variables occupied an intermediate position: sow_time = post (−0.115) was the most important management predictor (negative), confirming the

 doi:10.20944/preprints202603.1628.v1

detrimental effect of late sowing. Irrigation (Irr = Flow: +0.064) ranked 9th out of 24 in the absolute importance ranking, consistent with its role as a supplementary rather than primary driver of yield variability in the dataset.

3.5. Decision Tree Models: RandomTree and REPTree

The unpruned RandomTree grew to 398 nodes at depth 10 and performed on par with SMOreg (R = 0.903, RMSE = 0.194 t ha^{-1}, MAE = 0.112 t ha^{-1}), routing most of the predictive work through a clear hierarchy of soil and management splits. Its primary split was on *S_SILT* (threshold 19.5%), partitioning the dataset into low-silt soils (predominantly sandy-loam Entisols with lower PAW) and higher-silt soils (clay-loam Vertisols and Inceptisols). Within the high-silt branch, sowing date appeared at the second split level, and climate variables (ET_0, temperature, rainfall) from the third level onwards as refinement splits. Terminal node yields ranged from 3.95 t ha^{-1} (late sowing, rainfed, ET_0_dec_jan < 75 mm) to 6.42 t ha^{-1} (early sowing, S_SILT ≥ 43.5%). Supplemental irrigation consistently added 0.15-0.52 t ha^{-1} within all sowing-date terminal nodes.

Post-pruning condensed the tree architecture to just 73 nodes — roughly 18% of the RandomTree size — without meaningful loss of accuracy: R dropped by only 0.005 and RMSE increased by 0.004 t ha^{-1} (R = 0.898, RMSE = 0.198 t ha^{-1}, MAE = 0.139 t ha^{-1}), highlighting the compressibility of the AquaCrop response surface. Its primary split used *S_CLAY* (threshold 18%), with *sow_time* and *ET_0_dec_jan* as key secondary splitters. The difference in primary split variable between the two trees (S_CLAY vs. S_SILT) reflects the random attribute selection mechanism in RandomTree -both clay and silt are proxies of the same underlying soil water-retention gradient. Despite having only 18% of RandomTree's nodes, REPTree's RMSE is only 0.004 t ha^{-1} higher, demonstrating that post-pruning can substantially simplify the model without significant accuracy loss. The split-based variance-reduction importance for both tree models (Figure 3, bottom panels) corroborates the SMOreg weight vector: soil texture (S_SILT in RandomTree; S_CLAY in REPTree) accumulates the largest single importance share, followed by sowing date and the climate cluster, confirming that the predictor hierarchy is robust to the choice of tree algorithm. The observed-versus-predicted scatter plots (Figure 4) reveal a consistent tendency toward slight under-prediction at the highest yields (> 6.0 t ha^{-1}), corresponding to the most favourable soil–management combinations (fine-textured soils, standard sowing, supplemental irrigation), where the AquaCrop response surface exhibits mild non-linearities that tree-based splits cannot fully resolve. This upper-tail bias is small relative to the RMSE values in Table 1 and does not affect the direction of the agronomic signals.

3.6. Cross-Model Convergence of Agronomic Signals

Table 2 summarises the direction of the main predictor effects across all five models. All five algorithms consistently identified sowing date as the primary management discriminant: standard sowing (*reg*) > early sowing (*ant*) > late sowing (*post*) in LR, with consistent directional signals from SMOreg, and sow_time as the second-level split in both tree models. Supplemental irrigation carried a consistent positive signal in LR, SMOreg, and RandomTree. Soil texture (silt, clay) was positively associated with yield across all models where measurable, while sand exerted a negative influence. The convergence of these directional signals across fundamentally different algorithmic architectures -linear, kernel-based, and tree-based -provides robust, algorithm-independent support for the agronomic conclusions.

 doi:10.20944/preprints202603.1628.v1

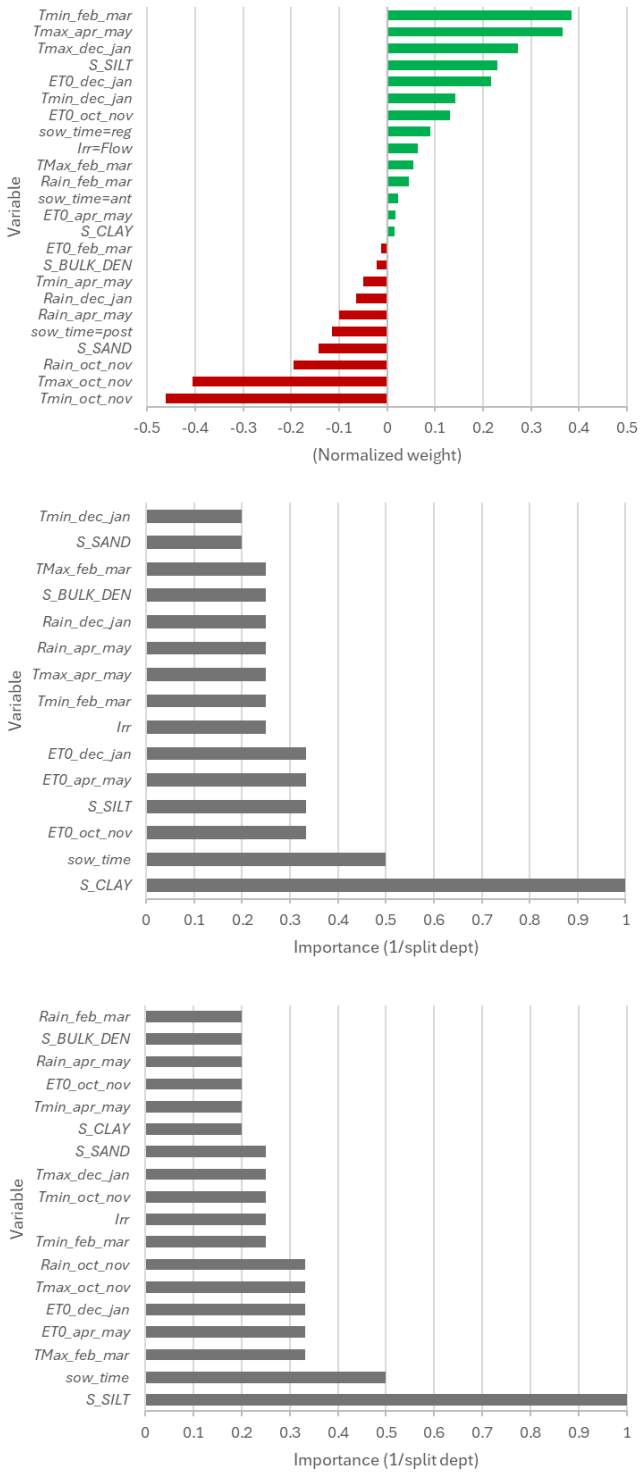

Figure 3. Normalised predictor importance across the three interpretable models. Top: SMOreg normalised weight vector; bottom centre and right: split-based variance-reduction importance for RandomTree and REPTree. Predictors are ranked by absolute importance; positive values indicate a yield-increasing effect, negative values a yield-decreasing effect.

 doi:10.20944/preprints202603.1628.v1

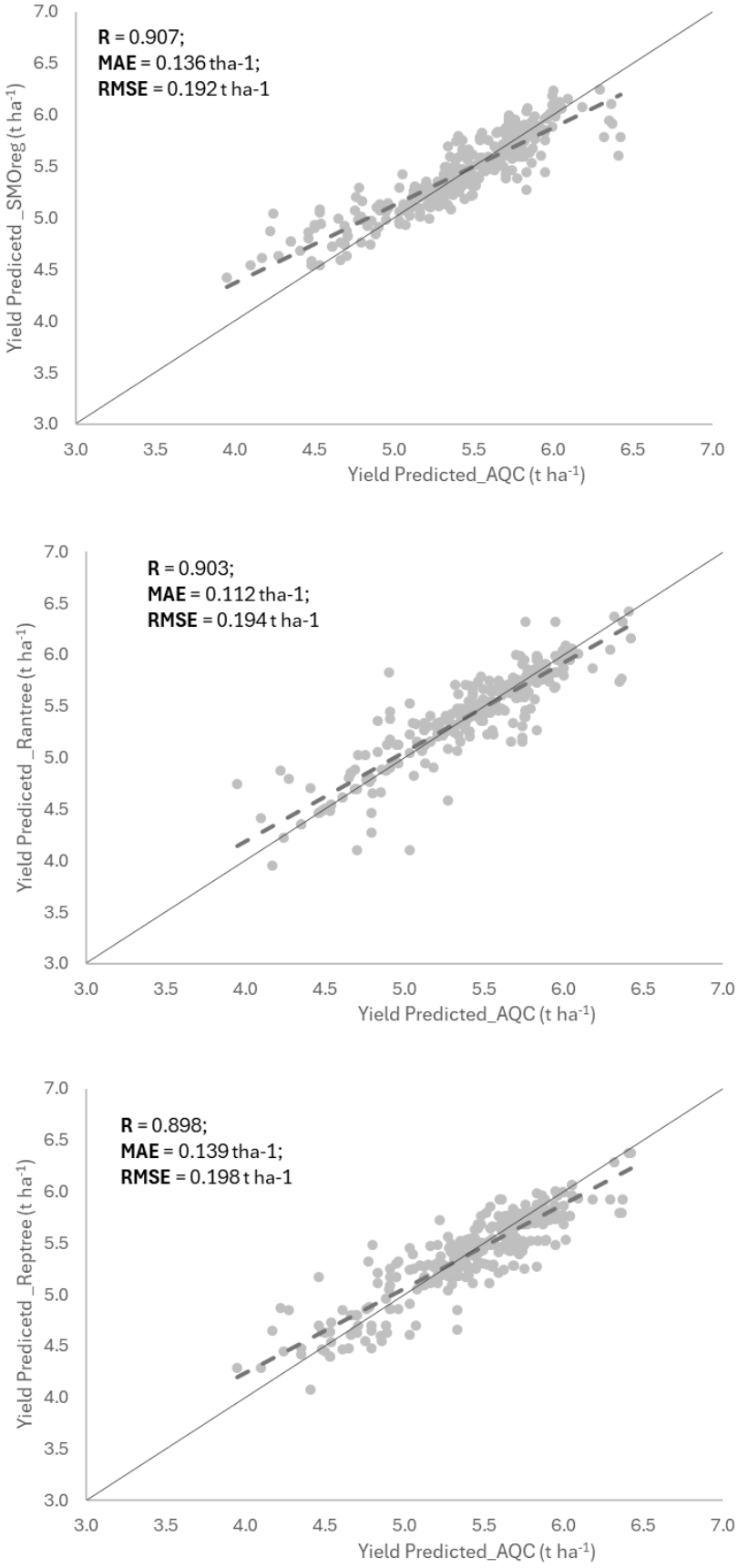

Figure 4. Observed versus predicted durum wheat grain yield (t ha^{-1}) for SMOreg, RandomTree, and REPTree from 10-fold cross-validation (n = 342). The dashed line is the 1:1 reference.

 doi:10.20944/preprints202603.1628.v1

Table 2. Cross-model summary of predictor effect directions on durum wheat yield. ++ strong positive; + positive; − negative; n.s. not retained by stepwise; ≈ near-zero; n.a. not attributable (MLP distributed weights); ref reference category.

Predictor	LR	MLP	SMOreg	RandTree	REPTree
Sow: standard (reg)	++	n.a.	+	✓	✓
Sow: early (ant)	+	n.a.	+	✓	✓
Sow: late (post)	ref	n.a.	−	✓	✓
Irrigation (Flow)	+	n.a.	+	+	≈
S_SILT	+	n.a.	+	✓ (1st split)	n.s.
S_CLAY	n.s.	n.a.	+	−	✓ (1st split)
S_SAND	−	n.a.	−	−	−
S_BULK_DEN	−	n.a.	≈	−	−
Tmin_oct_nov	≈	n.a.	−	−	−
Tmax_apr_may	≈	n.a.	+	−	−
Rain_apr_may	+	n.a.	−	+	+
Rain_oct_nov	−	n.a.	−	✓	−
ET_0_dec_jan	≈	n.a.	+	−	✓ (2nd split)

4. Discussion

The near-perfect accuracy of the MLP (R = 0.983, RMSE = 0.083 t ha^{-1}, approximately 3.4% of the observed yield range) is attributable to a mechanism well documented in the ML-crop model coupling literature: when training data are generated by a deterministic physical simulator rather than collected from noisy field measurements, the achievable accuracy is bounded only by the complexity of the model response surface and the capacity of the ML algorithm to represent it. Shahhosseini et al. [14] achieved R^2 up to 0.725 coupling ML with the APSIM crop model for US maize against observed county-level yields, while Gautam et al. [15] reported comparable findings comparing ANN and AquaCrop in Illinois — in both cases, the ML models outperformed the process-based benchmarks. The considerably higher accuracy obtained here for the MLP (R^2 = 0.966) is attributable to the deterministic, noise-free nature of AquaCrop simulation outputs used as training data — a mechanism well documented when ML algorithms are trained on simulation-derived rather than field-observed data, which eliminates stochastic measurement noise from the training signal [9,10]. Within this benchmark, the MLP result is consistent and expected; its primary value here is to establish an accuracy ceiling against which the interpretable models are evaluated.

The four interpretable algorithms — LR, SMOreg, RandomTree, and REPTree — clustered tightly at R = 0.891–0.907 (RMSE = 0.192–0.203 t ha^{-1}), separated by only 0.011 t ha^{-1}. In comparable field-based wheat yield studies such as Aravind et al. [29], R values of 0.70–0.85 and RMSE exceeding 0.5 t ha^{-1} are typical due to measurement uncertainty and spatial heterogeneity; therefore, clustering at R ≈ 0.90 represents a strong result for any single algorithm in this family. The narrow error band is consistent with the structured, low-noise nature of the training data, which reduces the variance component of prediction error for all algorithms equally. The quantified cost of full interpretability - ΔRMSE ≈ 0.115 t ha^{-1} relative to the MLP, corresponding to a 1.9% increase in error -has direct practical implications: the MLP is the optimal choice for high-precision spatial decision-support tools, while LR or REPTree are preferable for regulatory or advisory contexts where predictions must be transparent and defensible to farmers, advisors, or policy bodies [26]. This study provides a quantified basis for that choice in the durum wheat production context of the Capitanata plain.

 doi:10.20944/preprints202603.1628.v1

The most consequential result is the robust convergence of agronomic signals across all five algorithms. Sowing date emerged as the primary management discriminant: the LR estimated yield advantages of +0.531 t ha^{-1} for standard sowing and +0.380 t ha^{-1} for early sowing relative to late sowing; the SMOreg weight vector assigned the largest management weight to sow_time = post (−0.115); and both tree models placed sowing date at the second split level, immediately after the primary soil texture split. These estimates are consistent with independent field evidence from Southern Italy and comparable Mediterranean environments: Bassu et al. [16] demonstrated through long-term simulations in a Mediterranean clay soil environment that early sowing (October–November) provides yield advantages only in seasons with low waterlogging risk, while Ferrise et al. [31] showed that delayed sowing shortens the pre-anthesis vegetative phase and reduces grain-filling duration, with the associated higher pre-anthesis temperatures being a major determinant of final single grain dry matter, resulting in lower grain yields relative to autumn sowing. The slight advantage of standard over early sowing reflects the risk of excessive vegetative growth and frost damage in very early-sown crops under mild Mediterranean autumn conditions [30], a trade-off captured by the AquaCrop cold-stress and canopy cover routines. Soil texture occupied the top split position in both tree models (S_CLAY threshold 18% in REPTree; S_SILT threshold 19.5% in RandomTree) and ranked among the strongest predictors in SMOreg (S_SILT: +0.230; S_SAND: −0.143), while bulk density was the dominant soil predictor in LR (−0.618 t ha^{-1} per g cm^{-3}). Fine-textured, less compacted soils provide greater plant-available water capacity, deeper rooting, and better aeration [4,5], directly reducing the frequency and severity of water stress events in the AquaCrop soil-water balance and in turn improving simulated canopy expansion, biomass accumulation, and harvest index. That soil texture dominates the tree splits but ranks below sowing date in the LR coefficients reflects a methodological difference rather than a biological contradiction: trees optimise global variance reduction, maximised by the variable that most uniformly partitions the full data range, while LR estimates marginal effects holding all other variables constant [25]. Supplemental irrigation at anthesis exerted a consistent but moderate positive signal: +0.174 t ha^{-1} in LR, a positive SMOreg weight (+0.064), and a systematic 0.15-0.52 t ha^{-1} benefit across all RandomTree sowing-date terminal nodes. The agronomic basis is well established: anthesis is the most water-stress-sensitive phase of the wheat cycle, and even a mild transient deficit during pollination can cause spikelet sterility and reduce grains per spike [32]. Source-sink dynamics at this critical stage [33] mean that even a brief water deficit can cause irreversible reductions in grain number, underscoring the value of targeted irrigation at this phenological window. A supplemental application of approximately 50 mm effectively buffers this risk, as demonstrated in the AquaCrop simulations for the Capitanata plain [8]. The relatively modest magnitude of the irrigation signal compared to sowing date is consistent with the rainfed baseline of the Capitanata plain, where spring precipitation is adequate in most years and the marginal benefit of supplemental irrigation depends on the probability of transient water stress at flowering.

The SMOreg normalised weight vector revealed a strong dominance of early-season temperature over all other climate predictors. The two largest absolute weights were both negative autumn temperature terms -Tmin_oct_nov (−0.462) and Tmax_oct_nov (−0.405) -reflecting AquaCrop's phenological mechanism [4,5]: elevated cumulated thermal loads in October-November accelerate leaf area development and phenological progression of recently germinated seedlings, shortening the tillering phase and reducing root biomass accumulation at the onset of winter. The identical sign of both maximum and minimum temperature weights indicates that it is the overall thermal load -rather than frost risk -that governs this response. Conversely, mild late-winter nights (Tmin_feb_mar: +0.385) and high spring maximum temperature (Tmax_apr_may: +0.366) carried the largest positive climate weights, consistent with the agroclimatic optimum for durum wheat in Southern Italy: moderately cool autumn and winter for establishment and vernalisation, followed by a warm but not excessively hot spring for grain filling [30]. Bassu et al. [16], using long-term simulations in a Mediterranean clay soil environment, demonstrated that sowing date interacts with waterlogging risk to determine optimal management windows, with early sowing (October–

 doi:10.20944/preprints202603.1628.v1

November) providing yield advantages only in seasons or soils with low waterlogging risk. The secondary climate splits in the tree models -ET_0_dec_jan as a second-level splitter in REPTree, and autumn rainfall in RandomTree -broadly corroborate this seasonal hierarchy, strengthening the credibility of the climate interpretation through independent cross-algorithmic evidence.

The convergence of directional signals across five fundamentally different algorithmic architectures -linear, kernel-based, and tree-based -constitutes a form of algorithm-independent validation that reinforces the robustness of the agronomic conclusions beyond any single modelling assumption. Standard sowing, fine-textured soils, and moderately cool autumn temperatures consistently emerge as the primary drivers of yield advantage in the Capitanata plain, regardless of whether the prediction tool is a transparent regression equation, a kernel machine, or a pruned decision tree. This cross-algorithm convergence reflects the same principle emphasised in systematic reviews of ML in crop yield prediction [9,13]: the consistency of predictor importance rankings across methods provides stronger agronomic evidence than accuracy metrics alone, and constitutes a more reliable basis for management recommendations than any single algorithm, however accurate.

Several limitations of the present study should be acknowledged. The ML models were trained exclusively on AquaCrop simulation outputs and inherit the model's parameterisation assumptions: any systematic bias in the simulation of durum wheat under the specific soil, climate, and management conditions of the Capitanata plain would propagate directly to ML predictions, and external validation against independently observed field yield data remains a priority for future work. The AquaCrop parameter set was calibrated for a single representative modern cultivar [21], so cultivar-specific parameterisation would modify the quantitative estimates of sowing date and irrigation effects. Nitrogen dynamics and pest/disease pressure -which significantly modulate yield in real farming systems -are not included in the simulation framework and are therefore absent from the ML predictor space. Finally, the 10-year reference period (2014-2023) captures a specific segment of regional climate variability; extending simulations to a longer historical baseline would better represent the full distribution of climate anomalies and extreme events. Integration of the ML emulators with climate change projections [3] is a natural next step, though it requires running AquaCrop anew under future climate scenarios rather than extrapolating the current models beyond their training domain.

5. Conclusions

This study demonstrated that ML surrogates can efficiently emulate the AquaCrop-GIS response surface for durum wheat across the heterogeneous soils and climates of the Capitanata plain, with the MLP achieving near-perfect predictive accuracy ($R = 0.983$, RMSE = 0.083 t ha^{-1}) attributable to the deterministic, noise-free nature of simulation-derived training data. The four interpretable models -LR, SMOreg, RandomTree, and REPTree -clustered tightly at $R = 0.891$-0.907 (RMSE = 0.192-0.203 t ha^{-1}), separated by only 0.011 t ha^{-1}, confirming that the accuracy cost of choosing a transparent model is negligible for this class of structured training data. This quantified accuracy-interpretability trade-off provides a practical basis for tool selection: the MLP is optimal for spatial decision-support applications requiring high precision, whereas LR or REPTree are preferable where predictions must be explainable and defensible to farmers, advisors, or regulatory bodies.

Across all five algorithms, standard sowing (1 November) emerged as the dominant management driver, with a consistent yield advantage of +0.531 t ha^{-1} over late sowing (15 November), while early sowing (15 October) added +0.380 t ha^{-1}. Supplemental irrigation at anthesis contributed a stable +0.174 t ha^{-1} across all soil and climate combinations, confirming the critical sensitivity of grain set to transient water stress at flowering. Soil texture dominated the primary splits in both tree models and ranked among the top predictors in the SMOreg weight vector, with fine-textured soils (high silt, high clay) consistently producing higher yields through their greater plant-available water capacity. The SMOreg normalised weight vector further identified autumn temperature (Tmin_oct_nov, Tmax_oct_nov) as the leading climate signal, reflecting the AquaCrop phenological mechanism by which elevated early-season thermal loads curtail the tillering

 doi:10.20944/preprints202603.1628.v1

phase. The convergence of these directional signals across five fundamentally different algorithmic architectures -linear, kernel-based, and tree-based -provides robust, algorithm-independent validation of the key agronomic conclusions, and confirms that ML emulators trained on process-model outputs can deliver actionable, transparent guidance for durum wheat management in Mediterranean dryland farming systems.

Author Contributions: conceptualization P.G.; methodology, P.G.; software, P.G., A.R.B.C and M.R.; formal analysis P.G.; Writing – Original Draft preparation, P.G.; writing – review and editing, A.R.B.C. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. FAO. FAOSTAT. Crops and Livestock Products; Food and Agriculture Organization of the United Nations: Rome, Italy, 2023. Available online: https://www.fao.org/faostat (accessed on 1 January 2025).
2. ISTAT. Tavole Dati Indicatori Agro-Ambientali 2010–2020; Italian National Institute of Statistics: Rome, Italy, 2021.
3. Pörtner, H.-O.; Roberts, D.C.; Tignor, M.; Poloczanska, E.S.; Mintenbeck, K.; Alegría, A.; Craig, M.; Langsdorf, S.; Löschke, S.; Möller, V.; et al. (Eds.) Climate Change 2022: Impacts, Adaptation, and Vulnerability; Cambridge University Press: Cambridge, UK, 2022.
4. Steduto, P.; Hsiao, T.C.; Raes, D.; Fereres, E. AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agron. J. 2009, 101, 426–437. https://doi.org/10.2134/agronj2008.0139s
5. Raes, D.; Steduto, P.; Hsiao, T.C.; Fereres, E. AquaCrop—The FAO crop model to simulate yield response to water: II. Main algorithms and software description. Agron. J. 2009, 101, 438–447. https://doi.org/10.2134/agronj2008.0140s
6. Todorovic, M.; Albrizio, R.; Zivotic, L.; Saab, M.-T.A.; Stöckle, C.; Steduto, P. Assessment of AquaCrop, CropSyst, and WOFOST models in the simulation of sunflower growth under different water regimes. Agron. J. 2009, 101, 509–521. https://doi.org/10.2134/agronj2008.0166s
7. Katerji, N.; Campi, P.; Mastrorilli, M. Productivity, evapotranspiration, and water use efficiency of corn and tomato crops simulated by AquaCrop in the Mediterranean region. Agric. Water Manag. 2013, 130, 14–26. https://doi.org/10.1016/j.agwat.2013.08.007
8. Garofalo, P.; Cammerino, A.R.B. Modeling the performance of a continuous durum wheat cropping system in a Mediterranean environment: Carbon and water footprint at different sowing dates, under rainfed and irrigated water regimes. Agriculture 2025, 15, 259. https://doi.org/10.3390/agriculture15030259
9. van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. Comput. Electron. Agric. 2020, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709
10. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. Sensors 2018, 18, 2674. https://doi.org/10.3390/s18082674
11. Jeong, J.H.; Resop, J.P.; Mueller, N.D.; Fleisher, D.H.; Yun, K.; Butler, E.E.; Timlin, D.J.; Shim, K.-M.; Gerber, J.S.; Reddy, V.R.; Kim, S.-H. Random forests for global and regional crop yield predictions. PLOS ONE 2016, 11, e0156571. https://doi.org/10.1371/journal.pone.0156571
12. Filippi, P.; Jones, E.J.; Wimalathunge, N.S.; Somarathna, P.D.S.N.; Pozza, L.E.; Ugbaje, S.U.; Jephcott, T.G.; Paterson, S.E.; Whelan, B.M.; Bishop, T.F.A. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precis. Agric. 2019, 20, 1015–1029. https://doi.org/10.1007/s11119-019-09064-4

 doi:10.20944/preprints202603.1628.v1

13. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput. Electron. Agric. 2018, 151, 61–69. https://doi.org/10.1016/j.compag.2018.05.012
14. Shahhosseini, M.; Hu, G.; Huber, I.; Archontoulis, S.V. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci. Rep. 2021, 11, 1606. https://doi.org/10.1038/s41598-020-80820-1
15. Gautam, V.; Gani, A.; Pathak, S.; Shukla, A.K.; Varma, M. Evaluating crop yield prediction models using AquaCrop, semi-physical model and artificial neural networks. Sci. Rep. 2025, 15, 27494. https://doi.org/10.1038/s41598-025-13453-x
16. Bassu, S.; Asseng, S.; Motzo, R.; Giunta, F. Optimising sowing date of durum wheat in a variable Mediterranean environment. Field Crops Res. 2009, 111, 109–118. https://doi.org/10.1016/j.fcr.2008.11.002
17. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444. https://doi.org/10.1038/nature14539
18. Fernandez-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? J. Mach. Learn. Res. 2014, 15, 3133–3181.
19. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. Stat. Comput. 2004, 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88
20. Kayad, A.; Sozzi, M.; Gatto, S.; Whelan, B.; Sartori, L.; Marinello, F. Integration of artificial intelligence and remote sensing for crop yield prediction and crop growth parameter estimation in Mediterranean agroecosystems: Methodologies, emerging technologies, research gaps, and future directions. Eur. J. Agron. 2025, 163, 127894. https://doi.org/10.1016/j.eja.2025.127894
21. Trombetta, A.; Iacobellis, V.; Tarantino, E.; Gentile, F. Calibration of the AquaCrop model for winter wheat using MODIS LAI images. Agric. Water Manag. 2016, 164, 304–316. https://doi.org/10.1016/j.agwat.2015.10.013
22. Boogaard, H.; Schubert, J.; De Wit, A.; Lazebnik, J.; Hutjes, R.; Van der Grijn, G. Agri4Cast Resources Portal. CGMS-WOFOST Crop Growth Monitoring System Applied to Europe; European Commission, Joint Research Centre: Ispra, Italy, 2020. Available online: https://agri4cast.jrc.ec.europa.eu (accessed on 1 May 2024).
23. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements; FAO Irrigation and Drainage Paper 56; FAO: Rome, Italy, 1998.
24. FAO; IIASA. Harmonised World Soil Database Version 2.0; FAO: Rome, Italy; IIASA: Laxenburg, Austria, 2023; ISBN 978-92-5-137499-3.
25. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An Introduction to Statistical Learning; Springer: New York, NY, USA, 2013. https://doi.org/10.1007/978-1-0716-1418-1
26. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 2017, 30, 4765–4774.
27. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees; Wadsworth & Brooks/Cole: Belmont, CA, USA, 1984.
28. Zhao, Y.; Xiao, D.; Bai, H.; Tang, J.; Liu, D.L.; Qi, Y.; Shen, Y. The prediction of wheat yield in the North China Plain by coupling crop model with machine learning algorithms. Agriculture 2023, 13, 99. https://doi.org/10.3390/agriculture13010099
29. Aravind, K.S.; Vashisth, A.; Krishnan, P.; Das, B. Wheat yield prediction based on weather parameters using multiple linear, neural network and penalised regression models. J. Agrometeorol. 2022, 24, 18–25. https://doi.org/10.54386/jam.v24i1.1002
30. Flagella, Z.; Losavio, N.; Zatta, A.; Troccoli, A. Water relations and nitrogen uptake in durum wheat as affected by sowing date. Eur. J. Agron. 2002, 17, 221–230. https://doi.org/10.1016/S1161-0301(02)00005-8
31. Ferrise, R.; Triossi, A.; Stratonovitch, P.; Bindi, M.; Martre, P. Sowing date and nitrogen fertilisation effects on dry matter and nitrogen dynamics for durum wheat: An experimental and simulation study. Field Crops Res. 2010, 117, 245–257. https://doi.org/10.1016/j.fcr.2010.03.010

 doi:10.20944/preprints202603.1628.v1

32. Farooq, M.; Hussain, M.; Siddique, K.H.M. Drought stress in wheat during flowering and grain filling periods. Crit. Rev. Plant Sci. 2014, 33, 331–349. https://doi.org/10.1080/07352689.2014.875291
33. Slafer, G.A.; Savin, R. Source-sink relationships and grain mass at different positions within the spike in wheat. Field Crops Res. 1994, 37, 39–49. https://doi.org/10.1016/0378-4290(94)90080-9