

Within-orchard transfer of European hazelnut water-stress classification from a single reference station

Stanley Best^{1*}, José Maldonado¹, Cristóbal Campos², Jae Seung Noh³, and Francisco Altimiras⁴

¹Instituto de Investigaciones Agropecuarias, INIA Quilamapu, Chillán, Chile.

²Fundación Observatorio Ñuble, Chillán, Chile.

³National Institute of Agricultural Sciences, Rural Development Administration, Field Farming Mechanization Division, Jeonju-si, Republic of Korea.

⁴Universidad de Las Américas, Facultad de Ingeniería y Negocios, Santiago, Chile.

*Corresponding author (sbest@inia.cl).

Received: 11 August 2026; Accepted: 21 September 2026, Available online: 29 October 2026.

doi:10.4067/S0718-58392027000100091

ABSTRACT

Soil water content measured at a single depth or location may not represent irrigation need across a heterogeneous orchard. The objective was to determine whether a classification model developed at one reference station could identify water-stress states at three other stations of a commercial European hazelnut (*Corylus avellana* L.) 'Tonda di Giffoni' orchard without recalibration. Three management zones were delineated by k-means clustering of Sentinel-2 imagery acquired from 2023 to 2025. Station-specific irrigation thresholds (0.190 to 0.238 m³ m⁻³, all reached at a stem water-potential tension of 1.28 to 1.32 MPa) were defined from soil-water depletion and pressure-chamber measurements. An L2-penalized logistic model combined the reference-station Surface-to-Profile Water Extraction Ratio (SPWER), local PlanetScope canopy dynamics, and vapor pressure deficit. Development used 46 station-dates with 12 stress events in 18 irrigation cycles, and leave-one-cycle-out cross-validation yielded an area under the receiver operating characteristic curve (ROC-AUC) of 0.971. Applied without recalibration to 141 station-dates at three held-out stations, the model reached ROC-AUC 0.874 (95% confidence interval [CI] 0.814 to 0.928; lower limit above the 0.500 no-discrimination value), balanced accuracy 0.779, sensitivity 0.612, and specificity 0.946. A reduced model without SPWER reached ROC-AUC 0.849 and balanced accuracy 0.810, and a paired cycle bootstrap did not establish an incremental benefit of SPWER (difference 0.025; 95% CI -0.018 to 0.069). Twenty-three of 24 errors occurred within 0.02 m³ m⁻³ of the local threshold. The model separated within-orchard stress states, whereas the contribution of a single reference station was not resolved.

Key words: Irrigation management, management zones, precision agriculture, remote sensing, soil-plant-atmosphere continuum.

INTRODUCTION

Water scarcity is increasing the need for irrigation strategies that raise water productivity without compromising orchard performance (Jenkins and Block, 2024; Xing and Wang, 2024). Orchard water status results from soil water availability, root uptake, plant hydraulic regulation, and atmospheric demand, none of which is spatially uniform: Soil hydraulic properties, rooting depth, emitter performance, and canopy architecture vary over a few meters, so adjacent sectors differ in the water they store and in the rate at which trees deplete it (Ballester et al., 2024; Calabritto et al., 2024; Velázquez-Chávez et al., 2024). A uniform irrigation threshold is therefore difficult to justify, because the same volumetric water content may correspond to different plant responses in different sectors. Sensor placement guided by an explicit analysis of spatial

variability outperforms placement based on convenience (Dong et al., 2025; Best et al., 2026), and irrigation uniformity itself modifies the spatial structure that zoning attempts to capture (Bonzi et al., 2026).

Stem or trunk water potential remains the reference variable for interpreting local depletion, because it integrates the soil-plant-atmosphere continuum at the scale of the individual tree (Lakso et al., 2022; Blanco and Kalcsits, 2023; Conesa et al., 2023; Carella et al., 2024), but its measurement over space is laborious, which has motivated efforts to infer it from imagery in almond (Savchik et al., 2024) and olive (Garofalo et al., 2024). Canopy temperature can corroborate reduced transpirational cooling, although its interpretation depends on the atmospheric state at acquisition (Ahmad et al., 2021; Ramírez-Cuesta et al., 2022; Blanco and Kalcsits, 2024). Multitemporal Sentinel-2 imagery delineates zones that persist across seasons, whereas PlanetScope resolves individual crowns and describes canopy evolution within a cycle (Rouault et al., 2024). Because absolute vegetation indices respond to biomass, phenology, and soil background as well as to water status, normalization against a within-cycle baseline improves their interpretation (González-Dugo et al., 2021; Sandoñs-Pozo et al., 2022; Termin et al., 2023; González-Dugo, 2024).

The root zone provides a complementary signal. As the upper layers dry, uptake shifts towards deeper layers, and the vertical partitioning of extraction changes in a way that reflects fine-root distribution and local water content (Zhu et al., 2023). This compensation is imperfect and frequently overestimated by root-scale models under drying conditions (Khare et al., 2022), so an empirical descriptor of the partitioning may carry information a canopy-based approach does not. Stomatal closure in response to rising vapor pressure deficit is understood at the molecular level (Jalakas et al., 2021), and its sensitivity varies over the season (Guo et al., 2022).

A lower-cost architecture would combine one well-instrumented station as a temporal indicator of the irrigation cycle, satellite observations representing the spatial canopy response, and a meteorological record of atmospheric demand. Whether such a model can be moved to another position without recalibration is an open question. Transferability has been examined mainly for crop classification and yield prediction, where performance degrades outside the calibration domain by an amount not predictable from internal validation (Orynbaikyzy et al., 2022; Hoppe et al., 2024; Sahu et al., 2026). Comparable evidence for water-status classification within a single orchard is scarce.

The hypothesis was that stress classes defined independently at each station can be transferred within an orchard by combining reference-station root-zone extraction, local canopy evolution, and atmospheric demand. The objectives were to develop such a model at one station, evaluate it without recalibration at three stations not used in model construction, and determine by ablation whether the reference-station extraction ratio improves prediction beyond satellite and atmospheric variables.

MATERIALS AND METHODS

Study site

The study was conducted during the 2025-2026 irrigation season in a commercial European hazelnut (*Corylus avellana* L.) orchard at Las Nieves (36°39'23" S, 71°59'32" W; 198 m a.s.l.), Ñuble Region, Chile. The climate is Mediterranean with a dry summer, so essentially all crop water use between December and March must be supplied by irrigation. 'Tonda di Giffoni' trees, planted in 2016 at 4 × 6 m, were drip-irrigated with two laterals per row and emitters every 0.5 m (Rovira, 2021). The soil was an Andisol of the Arrayán series (Typic Melanoxerand), predominantly silt loam. In Andisols developed on volcanic ash the water actually available to roots varies more between positions than the textural class suggests, because pore structure is strongly modified by vegetation cover and organic fraction (Páez-Bimos et al., 2022). Published water-balance work in a drip-irrigated hazelnut orchard under a comparable Mediterranean semi-arid climate provides an independent reference for the seasonal water requirement (de la Fuente-Sáiz et al., 2026).

The orchard was managed commercially and no treatment was imposed. Replication was spatial rather than factorial: Four stations across three zones were followed through 18 to 19 irrigation cycles each, and every station-date constituted an observation.

Management zone delineation

Six cloud-free Sentinel-2 Level-2A images (10 m spatial resolution) acquired in the second half of January in 2023, 2024, and 2025 represented periods of high crop water demand; restricting the stack to one phenological

window in three seasons captures the persistent component of the spatial pattern. After cloud and shadow masking, the normalized difference vegetation index (NDVI) was calculated for each date, and the stack was clustered by k-means with the Cluster Analysis plugin of QGIS 3.34 (QGIS Development Team, Grüt, Switzerland). Solutions from $k = 1$ to 10 were compared by the within-cluster sum of squares (WCSS); the principal elbow occurred at $k = 3$ and the three zones were spatially contiguous, the property required for them to serve as sampling strata (Figure 1).

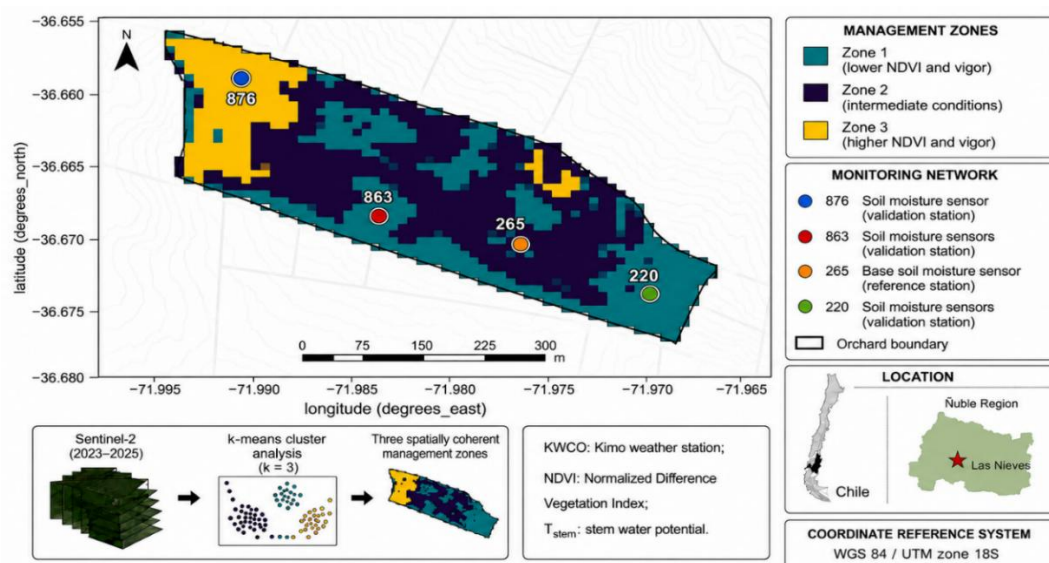


Figure 1. Management zones and monitoring network in the European hazelnut orchard at Las Nieves, Ñuble Region, Chile. Zones were delineated by k-means clustering ($k = 3$) of six Sentinel-2 images acquired in the second half of January from 2023 to 2025, and were used as sampling strata. Station 265, in zone 2, provided the data for model development; stations 220 and 876, in zone 1, and station 863, in zone 3, were held out for validation. The inset shows the location of the site within the Ñuble Region. NDVI: Normalized difference vegetation index.

Monitoring stations and irrigation thresholds

Reference station 265 (GeoTrace SpA, Chillán, Chile) was located in management zone 2, station 863 in zone 3, and stations 220 and 876 at two positions in zone 1 (Figure 1); placing two stations in one zone tests whether a remotely sensed zone is internally homogeneous with respect to irrigation need. Each station carried capacitance soil water content probes (GeoTrace SpA, Chillán, Chile) at 30 and 50 cm depth and an irrigation-line pressure sensor. A weather station (KNCO Co. Ltd., Seoul, Republic of Korea) recorded air temperature and relative humidity, from which vapor pressure deficit (VPD) was calculated. Soil water content at 15:00 h was synchronized with the weather and satellite records.

During progressive drying in January and February 2026, stem water potential was measured near solar noon on one instrumented tree per station. Three mature, sun-exposed leaves per tree were bagged in aluminized envelopes for at least 1 h to equilibrate with the stem and measured with a Scholander-type pressure chamber (PMS-1000; PMS Instrument Company, Albany, Oregon, USA); their mean represented the station-date. Canopy temperature was a qualitative field check only and was not analyzed quantitatively.

Stem water potential is reported as the absolute value of pressure-chamber tension. For each station, the operational irrigation threshold was the soil water content at 15:00 h at the first persistent departure from the preceding post-irrigation depletion trend, where the mean stem water-potential tension reached 12.8 to 13.2 bar. The rationale is hydraulic: While drainage governs depletion, water content falls approximately linearly; once uptake becomes limited by the conductance of the drying soil and the trees regulate stomatally, the curve bends, marking the transition from a soil-controlled to a plant-controlled regime (Figure 2). The same rule was applied at all four stations, so any difference between thresholds reflects the site and not the criterion (Table 1).

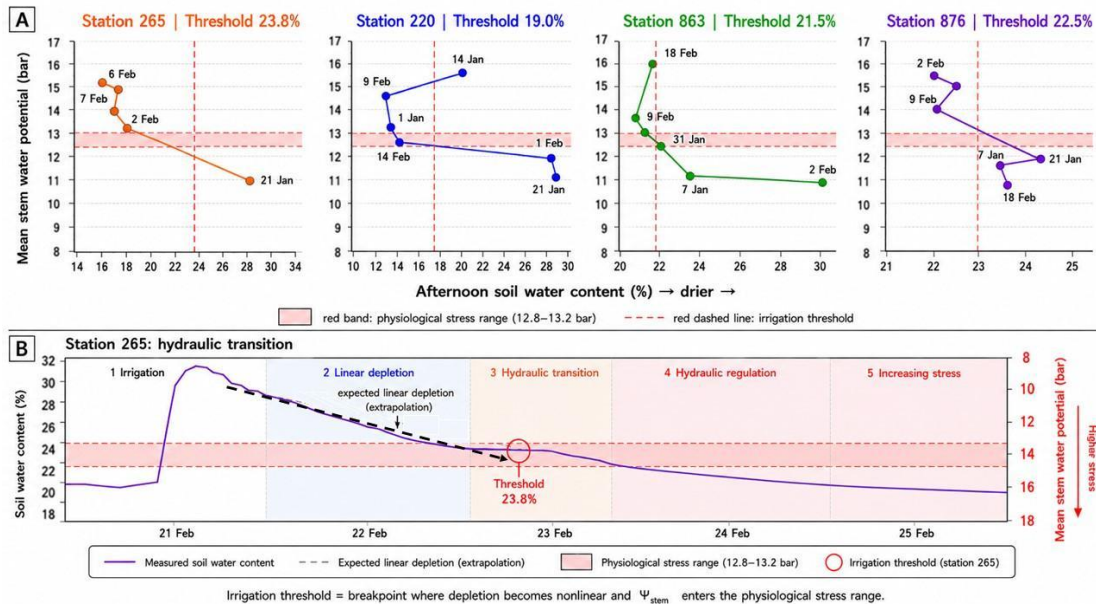


Figure 2. Physiological characterization of the monitoring stations and definition of the operational irrigation thresholds. (A) Relationship between afternoon soil water content and mean stem water-potential tension at each station; the vertical line marks the station-specific threshold and the shaded band the stress range of 1.28 to 1.32 MPa. (B) Soil-water depletion sequence at reference station 265, in which the threshold corresponds to the departure from the linear phase governed by gravitational drainage towards the non-linear phase governed by plant hydraulic regulation. Stem water potential is expressed as the absolute value of pressure-chamber tension.

Table 1. Station-level evidence and operational irrigation thresholds in a European hazelnut orchard at Las Nieves, Ñuble Region, Chile, during the 2025-2026 season. Thresholds were defined at a stem water-potential tension of 1.28 to 1.32 MPa by the persistent departure from the post-irrigation depletion trend, applying the same graphical rule at every station. Each station-date is the mean of three sun-exposed leaves on one instrumented tree. Stem water potential is the absolute value of pressure-chamber tension; canopy temperature served only as a qualitative field check.

Station	Management zone/Role	Irrigation threshold ($\text{m}^3 \text{m}^{-3}$)
265	2/Development	0.238
220	1/Validation	0.190
863	3/Validation	0.215
876	1/Validation	0.225

PlanetScope observations and predictors

Analysis-Ready PlanetScope Surface Reflectance imagery (ARPS-SR; 3 m spatial resolution) was evaluated from 2 December 2025 to 1 March 2026. The product is orthorectified, atmospherically corrected, and harmonized to the Sentinel-2 spectral bands, and its quality layers were used for cloud and shadow screening (Planet Labs PBC, 2026). Harmonization matters because the zones derive from Sentinel-2 and the canopy predictors from PlanetScope. A crown mask excluded inter-row vegetation and bare soil. Fifty-one dates contributed usable observations from 16 December to 25 February; missing images were not interpolated, and an interrupted baseline-to-follow-up sequence excluded the current cycle (Table 2).

Table 2. PlanetScope record availability and retained irrigation cycles per monitoring station. A cycle began at a recorded irrigation event; a gap longer than 4 d after missing imagery initiated a new retained cycle. All records of an omitted cycle were withheld together in cross-validation.

Station	Management zone/Role	Station-dates	Stress/Non-stress	Retained interval	Irrigation cycles
265	2/Development	46	12/34	16 Dec-25 Feb	18
863	3/Validation	46	12/34	16 Dec-25 Feb	19
876	1/Validation	48	24/24	16 Dec-23 Feb	19
220	1/Validation	47	13/34	16 Dec-25 Feb	17

Irrigation onset was identified from line pressure together with the increase in soil water content, and the first image after onset defined the cycle baseline. The Surface-to-Profile Water Extraction Ratio was calculated as $SPWER = E_{30}/(E_{30} + E_{50})$, where E_{30} and E_{50} were the daily reductions in soil water content at 30 and 50 cm depth. The ratio is dimensionless and bounded between 0 and 1; values near 1 indicate extraction concentrated in the shallow layer, as expected soon after irrigation, whereas declining values indicate a shift towards the deeper layer as the surface dries. The canopy persistence index (CPI) was the daily percentage change in crown NDVI relative to the cycle baseline, $CPI = 100[(NDVI_t - NDVI_{0,k})/NDVI_{0,k}]/dt$, referred to the baseline image of cycle k and reset at every irrigation event, and the cumulative canopy response (CCR) accumulated the sign of the daily canopy change (+1, -1, or 0) within the cycle, so CPI describes the instantaneous rate and CCR the persistence of the trajectory. Vapor pressure deficit was calculated from air temperature (T , in °C) and relative humidity (RH, in %) as $VPD = e_s(T)[1 - RH/100]$, with $e_s(T) = 0.6108 \exp[17.27T/(T + 237.3)]$ in kPa.

Outcome definition and datasets

Each station-date was classified as stressed when its soil water content at 15:00 h was equal to or lower than the station-specific threshold. Model development used station 265 alone ($n = 46$; 12 stress and 34 non-stress station-dates in 18 cycles). Held-out validation used stations 220, 863, and 876 ($n = 141$; 49 stress and 92 non-stress station-dates in 55 cycles; Table 2). Validation predictors comprised SPWER from station 265, the canopy indices of the local crown, and orchard VPD; local soil water at the validation stations defined the observed class only and never entered the model, so those stations played no part in threshold selection, variable definition, preprocessing, or fitting. The 18 development cycles contributed one to four retained station-dates each, and the 12 stress observations were distributed over eight of them, so the stress events were not concentrated in a single cross-validation fold.

Statistical analysis

Each predictor was summarized by median and interquartile range within the two outcome classes; Shapiro-Wilk tests assessed the marginal distributions and Spearman rank correlations the association among predictors. These summaries were descriptive; variable inclusion was fixed a priori by the conceptual model.

An L2-penalized binary logistic regression was fitted because the development sample contained only 12 stress station-dates, a ratio of events to predictors at which an unpenalized fit would give unstable coefficients. Missing predictors were replaced by development-set medians and standardized with development-set means and standard deviations, both recomputed inside each cross-validation fold to avoid leaking information from the withheld cycle. The penalty parameter was $C = 1.0$. Analyses used Python 3.12 (Python Software Foundation, Beaverton, Oregon, USA) and scikit-learn 1.8.0 (global community of open-source) with the lbfgs solver and 10 000 iterations, under which the base fit is deterministic; seed 20260728 was used only for bootstrap resampling.

Internal performance was estimated by leave-one-irrigation-cycle-out cross-validation (LOCO-CV). Withholding the complete cycle is essential because observations within a cycle share the same baseline image and depletion sequence and are not independent; a record-wise split would leak information and inflate the estimate. The fitted full model was then applied without recalibration to the three held-out stations at a probability threshold of 0.50. Uncertainty was estimated with 2000 bootstrap resamples of complete station-specific cycles, with 95% confidence intervals from the 2.5th and 97.5th percentiles.

To test the incremental contribution of the reference sensor, the full model (SPWER, CPI, CCR, and VPD) was compared with a reduced model (CPI, CCR, and VPD) under identical preprocessing, cycle cross-validation, and held-out transfer. Paired bootstrap differences were obtained by evaluating both fitted models on the same resampled validation cycles, which removes the resampling variance common to both. A difference was regarded as supported only when its 95% confidence interval excluded zero. Discrimination was assessed by the areas under the receiver operating characteristic and precision-recall curves, calibration by the Brier score with calibration intercept and slope, and classification by sensitivity, specificity, and balanced accuracy.

Spatial application

For each retained image date, the SPWER of the reference station and orchard VPD were combined with crown-level CPI and CCR to calculate full-model stress probabilities, and crowns with a probability of 0.50 or greater were displayed as stressed (Figure 3). The map applies the station-validated model to positions not independently observed, and is a screening output rather than a validated prescription.

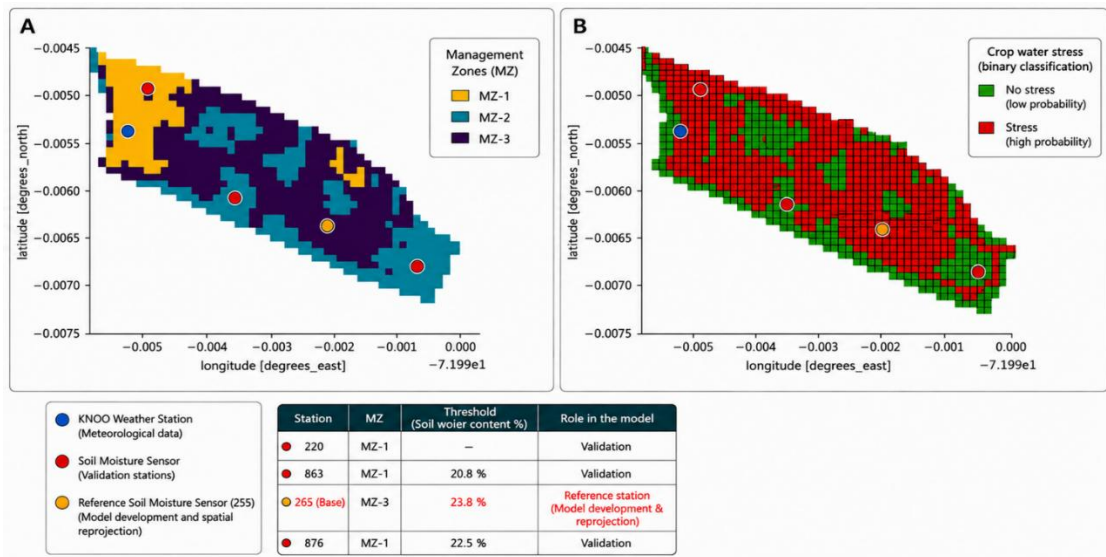


Figure 3. Spatial application of the full logistic model to the masked orchard crowns for a single PlanetScope image date. Crown units with a stress probability equal to or greater than 0.50 are displayed as stressed. Independent validation was limited to stations 220, 863, and 876; the continuous surface is an extrapolated model output and not an independently validated irrigation prescription. SPWER: Surface-to-profile water extraction ratio; CPI: canopy persistence index; CCR: cumulative canopy response; VPD: vapor pressure deficit.

RESULTS AND DISCUSSION

Management zones and station-specific irrigation thresholds

The within-cluster sum of squares (WCSS) curve showed its principal elbow at $k = 3$, and the selected solution delineated three contiguous zones that guided the four-station design (Figure 1). At a common stem water-potential tension of 1.28 to 1.32 MPa, the station-specific thresholds ranged from 0.190 to 0.238 $\text{m}^3 \text{m}^{-3}$ (Table 1; Figure 2). The two stations in management zone 1 had criteria of 0.190 and 0.225 $\text{m}^3 \text{m}^{-3}$, a difference of 0.035 $\text{m}^3 \text{m}^{-3}$ within a single zone, and the full range of 0.048 $\text{m}^3 \text{m}^{-3}$ means that a single orchard-wide criterion would have classified the monitored sectors differently at an equivalent plant water status. Andisols store a large amount of water, but the accessible fraction depends on effective depth, pore architecture, and fine-root distribution, all of which vary at short distances in soils derived from volcanic ash (Páez-Bimos et al., 2022), consistent with reports of contrasting soil-moisture thresholds among positions of the same orchard (Ballester et al., 2024; Calabritto et al., 2024).

Zones derived from multitemporal normalized difference vegetation index (NDVI) capture the persistent component of canopy vigor, which integrates many factors besides water availability, and the delineation is further influenced by irrigation uniformity (Bonzi et al., 2026). A remotely sensed zone should therefore be treated as a stratum for allocating measurement effort, not as a unit within which a single threshold can be assumed, which supports placing sensors by explicit analysis of spatial variability (Dong et al., 2025) and revising boundaries as the season progresses (Termin et al., 2023).

Behavior of the predictors

Surface-to-profile water extraction ratio (SPWER) was markedly lower in stress than in non-stress station-dates (median 0.293, interquartile range 0.263 to 0.324, against 0.783 and 0.391 to 1.000), as were canopy persistence index (CPI) (-2.0 and -3.0 to -2.0, against 0.0 and -1.0 to 1.0) and cumulative canopy response (CCR) (-2.746 and -3.989 to -1.908, against 0.000 and -1.623 to 1.956). The vapor pressure deficit (VPD) distributions overlapped substantially (3.115 and 2.766 to 3.497 kPa, against 2.911 and 2.550 to 3.425 kPa). The SPWER, CPI, and CCR departed from normality in both outcome classes (Shapiro-Wilk $W = 0.732$ to 0.914 ; $p \leq 0.0193$), whereas VPD did not ($W = 0.971$ and 0.974 ; $p = 0.4919$ and 0.9495), and CPI and CCR were strongly correlated (Spearman $\rho = 0.861$) while their correlations with VPD were weak.

These patterns are physiologically coherent. Stress station-dates occurred late in the depletion sequences, when extraction was no longer concentrated in the shallow layer and the canopy indices had already declined, the signature expected when uptake shifts towards deeper soil (Zhu et al., 2023). Atmospheric demand alone does not separate the classes, because the same evaporative demand produces different responses depending on root-zone water availability, and stomatal sensitivity to VPD varies over the season (Jalakas et al., 2021; Guo et al., 2022). The near-zero adjusted CCR coefficient indicates an absence of conditional information in a small dataset, not that cumulative canopy response lacks biological meaning.

Model fit and transfer to held-out stations

The fitted full-model equation was $\text{logit}(p) = -2.3319 - 1.5054 z(\text{SPWER}) - 1.6491 z(\text{CPI}) + 0.0224 z(\text{CCR}) + 0.2187 z(\text{VPD})$, where z denotes standardization with the development-set statistics. The signs are those expected from the conceptual model: The probability of stress increases as extraction shifts away from the shallow layer and as the canopy trajectory declines, and increases slightly with atmospheric demand. Expressed as odds ratios per standard deviation, the effects were 0.222 for SPWER (95% confidence interval [CI] 0.147 to 0.364) and 0.192 for CPI (0.100 to 0.368), whereas the intervals of CCR (1.023; 0.549 to 1.633) and VPD (1.244; 0.744 to 2.086) included 1, so only the extraction ratio and the canopy rate had individually resolved effects (Table 3). Cycle cross-validation at the reference station yielded ROC-AUC 0.971 and balanced accuracy 0.944.

Table 3. Standardized coefficients and bootstrap odds ratios (OR) of the full logistic model fitted at reference station 265. SD: Standard deviation; CI: confidence interval; SPWER: surface-to-profile water extraction ratio; CPI: canopy persistence index; CCR: cumulative canopy response; VPD: vapor pressure deficit. Coefficients were standardized with the development-set statistics and the intercept was -2.3319. Percentile intervals come from 2000 resamples of complete cycles; those of CCR and VPD include 1.

Term	Coefficient	OR per 1 SD	Bootstrap 95% CI
SPWER	-1.505	0.222	0.147 to 0.364
CPI	-1.649	0.192	0.100 to 0.368
CCR	0.022	1.023	0.549 to 1.633
VPD	0.219	1.244	0.744 to 2.086

Applied without recalibration, the model transferred to the three held-out stations with ROC-AUC 0.874 (95% CI 0.814 to 0.928), balanced accuracy 0.779, sensitivity 0.612, specificity 0.946, and Brier score 0.137 (Table 4). The confusion matrix contained 30 true positives, 87 true negatives, 19 false negatives, and five false positives. The lower confidence limit of 0.814 lies well above the 0.500 expected in the absence of discrimination, so the ranking of stress risk across positions is supported by the data. Average precision was

0.827 (95% CI 0.730 to 0.907) against a held-out stress prevalence of 0.348, and at the 0.50 operating point precision reached 0.857 (0.730 to 0.966) while sensitivity remained at 0.612 (0.476 to 0.745), so station-dates declared stressed were reliable whereas stress dates close to the local threshold were frequently missed. Calibration was acceptable although slightly overconfident at the extremes (intercept 0.542, slope 0.694, expected calibration error 0.086; Table 5).

Table 4. Performance of the full and reduced logistic models in leave-one-cycle-out cross-validation (LOCO-CV) at the reference station and in held-out within-orchard transfer. ROC-AUC: Area under the receiver operating characteristic curve; SPWER: surface-to-profile water extraction ratio; CPI: canopy persistence index; CCR: cumulative canopy response; VPD: vapor pressure deficit. The 95% confidence interval of the held-out ROC-AUC of the full model was 0.814 to 0.928.

Model	Dataset	ROC-AUC	Balanced accuracy	Sensitivity	Specificity	Brier score
Full: SPWER + CPI + CCR + VPD	Reference LOCO-CV	0.971	0.944	0.917	0.971	0.061
Reduced: CPI + CCR + VPD	Reference LOCO-CV	0.929	0.944	0.917	0.971	0.078
Full: SPWER + CPI + CCR + VPD	Held-out stations	0.874	0.779	0.612	0.946	0.137
Reduced: CPI + CCR + VPD	Held-out stations	0.849	0.810	0.673	0.946	0.132

Table 5. Precision-recall and calibration diagnostics of the full model in held-out within-orchard transfer. CI: Confidence interval; PR-AUC: area under the precision-recall curve. Intervals were obtained from 2000 resamples of complete held-out station cycles; calibration was assessed with five equal-frequency probability groups.

Metric	Estimate	Cycle-bootstrap 95% CI
Average precision	0.827	0.730 to 0.907
PR-AUC	0.825	0.723 to 0.905
Precision at $p > 0.50$	0.857	0.730 to 0.966
Sensitivity at $p > 0.50$	0.612	0.476 to 0.745
Brier score	0.137	0.100 to 0.180
Stress prevalence	0.348	-
Calibration intercept	0.542	-
Calibration slope	0.694	-
Expected calibration error (5 bins)	0.086	-

The loss of about 0.10 in ROC-AUC between internal cross-validation and external transfer is the central quantitative result. Comparable degradation on moving a model outside its calibration domain has been documented repeatedly for remotely sensed agricultural models, and the consistent lesson is that internal validation does not predict the magnitude of the loss (Orynbaikyzy et al., 2022; Hoppe et al., 2024; Sahu et al., 2026). The high specificity means few non-stress dates were flagged, an attractive property for a screening tool, whereas the sensitivity of 0.612 means a substantial proportion of genuine stress dates were missed, which precludes direct use as an irrigation trigger.

Incremental contribution of the reference-station sensor

The reduced model had a lower internal ROC-AUC than the full model (0.929 against 0.971), yet both produced the same cycle-cross-validated classification at the 0.50 threshold. In held-out transfer it reached ROC-AUC 0.849, balanced accuracy 0.810, sensitivity 0.673, specificity 0.946, and Brier score 0.132 (Table 4). Adding SPWER increased ROC-AUC by 0.025, but the paired cycle-bootstrap interval included zero (95% CI -0.018 to 0.069); the differences in balanced accuracy (-0.031; 95% CI -0.083 to 0.015) and Brier score (0.005; 95% CI -0.013 to 0.025) likewise gave no support to a consistent improvement, and no difference in discrimination, classification, or calibration excluded zero (Table 6).

Table 6. Paired difference between the full and the reduced logistic model in held-out within-orchard transfer. CI: Confidence interval; ROC-AUC: area under the receiver operating characteristic curve. The reduced model contained canopy persistence index (CPI), cumulative canopy response (CCR), vapor pressure deficit (VPD), and the full model added surface-to-profile water extraction ratio (SPWER). Positive differences favor the full model, except for the Brier score, where lower values are better. Both models were fitted once to the same development data and evaluated on the same 2000 resamples of complete held-out station cycles (seed 20260728).

Metric	Difference	Cycle-bootstrap 95% CI
ROC-AUC	0.025	-0.018 to 0.069
Average precision	0.002	-0.052 to 0.049
Accuracy	-0.021	-0.060 to 0.014
Balanced accuracy	-0.031	-0.083 to 0.015
Sensitivity	-0.061	-0.167 to 0.021
Specificity	0.000	-0.033 to 0.031
F1	-0.044	-0.125 to 0.022
Brier score	0.005	-0.013 to 0.025

This is a negative result and is reported as such. The SPWER was conceived as a temporal hydraulic anchor indicating where the orchard stands within its depletion cycle, not as a measurement of water status at every crown. The ablation shows the canopy and atmospheric variables already carried most of that information, and the reduced model recovered three additional stress station-dates at the same specificity. The evidence supports the transfer framework more strongly than it supports SPWER as a necessary component.

A mechanistic reading suggests when the variable should help. A single reference extraction signal can represent the whole orchard only while irrigation timing and root-zone depletion remain approximately synchronized across positions. Differences in soil depth, texture, emitter discharge, or rooting density shift the phase of the depletion curve between sectors, and once the phases diverge the reference signal describes a state the remaining sectors have not reached or have already passed. Compensatory uptake from deeper layers is itself imperfect and commonly overestimated under drying conditions (Khare et al., 2022). The hypothesis that follows is testable: The incremental value of a reference extraction signal should increase with irrigation synchrony and decrease with soil heterogeneity.

Distribution of classification errors

Twenty-three of the 24 errors of the full model occurred within $0.02 \text{ m}^3 \text{ m}^{-3}$ of the local threshold. All 19 false negatives lay 0.0001 to $0.0123 \text{ m}^3 \text{ m}^{-3}$ below it, and four of the five false positives lay 0.0052 to $0.0199 \text{ m}^3 \text{ m}^{-3}$ above it; the remaining false positive lay $0.0895 \text{ m}^3 \text{ m}^{-3}$ above the criterion.

This concentration is the most operationally relevant finding. The model is not confusing wet sectors with dry ones; it disagrees with a binary label precisely where that label discretizes a continuous process, and where the observed class is most sensitive to probe measurement error. The single error far from the threshold warrants inspection of that station-date. Read this way, the interval above the threshold behaves as a pre-stress or field-inspection band rather than a zone of model failure. This interpretation is post hoc, was not prespecified, and does not constitute a validated third class; it is offered as a hypothesis for the next evaluation, in which such a band should be defined in advance and tested.

Spatial application and limitations

Figure 3 shows the model applied across the masked crowns for a single image date. The classified surface is spatially structured rather than random and broadly consistent with the management zones, the behavior expected if the canopy predictors respond to the same variation in soil water supply. The map rests on an extrapolation: Independent observations exist only at the three held-out stations. Its legitimate use is to rank sectors for field inspection; driving valve operation would require validation at positions sampled across the map.

The study was restricted to one orchard and one irrigation season. Development rested on 46 station-dates and 12 stress events, with one instrumented tree per station, so biological replication is minimal and the

thresholds carry uncertainty not propagated into the classification. The held-out stations shared dates, weather, cultivar, soil series, and management with the reference station, which makes this a test of spatial transfer within an orchard rather than transfer to another orchard or season. Further evaluation should lock the model before testing complete seasons and independent properties, increase replication within zones, and include soils and irrigation systems with contrasting depletion timing. Prespecified comparisons should assess whether actual evapotranspiration from energy-balance approaches such as METRIC, thermal anomalies, soil properties, and irrigation-delivery variables improve calibration and sensitivity (Allen et al., 2007; Ahmad et al., 2021; Ramírez-Cuesta et al., 2022; Nevo et al., 2025), and embedded microtensiometers provide temporal resolution that pressure-chamber campaigns cannot achieve (Lakso et al., 2022; Savchik et al., 2024).

CONCLUSIONS

A logistic model developed at one reference station discriminated the station-specific water-stress classes of three other locations within the same orchard, which supports the hypothesis that stress states can be transferred within an orchard from a single instrumented position. In held-out transfer the model reached an area under the receiver operating characteristic curve of 0.874, with a lower confidence limit of 0.814 that lies above the value expected in the absence of discrimination, and a specificity of 0.946; its sensitivity was 0.612, a value insufficient for automatic irrigation triggering.

The irrigation thresholds required to reach an equivalent plant water status differed by 0.048 cubic meter per cubic meter among the four stations, and by 0.035 cubic meter per cubic meter between two stations located in the same management zone, so a single orchard-wide soil water criterion would have classified these sectors differently. Zones delineated from satellite imagery served as sampling strata but did not define units of uniform irrigation need.

Adding the extraction ratio of the reference station increased the area under the curve by 0.025 relative to a model based on satellite and atmospheric variables alone; the corresponding confidence interval included the absence of improvement and the reduced model had higher sensitivity, so the data do not demonstrate that the reference-sensor variable contributes predictive value. Twenty-three of the 24 classification errors occurred within $0.02 \text{ m}^3 \text{ m}^{-3}$ of the local irrigation threshold, showing that the disagreements concentrated where the binary criterion divides a continuous depletion process.

Author contribution

Conceptualization: S.B. Methodology: S.B., J.M. Software: C.C. Validation: S.B., J.M. Formal analysis: S.B., C.C. Investigation: S.B., J.M., F.A. Resources: S.B., J.M., J.S.N. Data curation: S.B., C.C. Writing-original draft: S.B. Writing-review & editing: S.B., J.M., C.C., J.S.N., F.A. Visualization: S.B., C.C., F.A. Supervision: S.B. Project administration: S.B., J.S.N. Funding acquisition: S.B., J.S.N. All co-authors reviewed the final version and approved the manuscript before submission.

Acknowledgements

This research was funded by the Instituto de Investigaciones Agropecuarias, Chile, and the Rural Development Administration, Republic of Korea, through the bilateral project “Enhancement of smart irrigation technology leveraging satellite and crop data for application across various crops”.

References

- Ahmad, U., Alvino, A., Marino, S. 2021. A review of crop water stress assessment using remote sensing. *Remote Sensing* 13:4155. doi:10.3390/rs13204155.
- Allen, R.G., Tasumi, M., Trezza, R. 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-Model. *Journal of Irrigation and Drainage Engineering* 133:380-394. doi:10.1061/(ASCE)0733-9437(2007)133:4(380).
- Ballester, C., Filev-Maia, R., Hornbuckle, J. 2024. Impact of soil spatial variability on young almond trees: A case study on heavy clay soils. *Agrosystems, Geosciences & Environment* 7:e20572. doi:10.1002/agg2.20572.
- Best, S., Maldonado, J., Navarro, M., De la Barrera, V. 2026. Digital tools in the context of Agriculture 4.0 for integrated management of orchard crops. In Saiz-Rubio, V., Rovira-Mas, F., Kole, C. (eds.) *Smart agriculture: Concepts, strategies, and case studies*. CRC Press, Boca Raton, Florida, USA.
- Blanco, V., Kalcsits, L. 2023. Long-term validation of continuous measurements of trunk water potential and trunk diameter indicate different diurnal patterns for pear under water limitations. *Agricultural Water Management* 281:108257. doi:10.1016/j.agwat.2023.108257.

- Blanco, V., Kalcsits, L. 2024. Relating microtensiometer-based trunk water potential with sap flow, canopy temperature, and trunk and fruit diameter variations for irrigated 'Honeycrisp' apple. *Frontiers in Plant Science* 15:1393028. doi:10.3389/fpls.2024.1393028.
- Bonzi, L., Carrara, M., Hamouda, F., Marasco, S., Massai, R., Puig-Sirera, À., et al. 2026. Effect of irrigation uniformity on the delineation of homogeneous zones in a pear orchard. *Agricultural Water Management* 326:110224. doi:10.1016/j.agwat.2026.110224.
- Calabritto, M., Mininni, A.N., Di Biase, R., Pietrafesa, A., Dichio, B. 2024. Spatio-temporal dynamics of root water uptake and identification of soil moisture thresholds for precision irrigation in a Mediterranean yellow-fleshed kiwifruit orchard. *Frontiers in Plant Science* 15:1472093. doi:10.3389/fpls.2024.1472093.
- Carella, A., Bulacio Fischer, P.T., Massenti, R., Lo Bianco, R. 2024. Continuous plant-based and remote sensing for determination of fruit tree water status. *Horticulturae* 10:516. doi:10.3390/horticulturae10050516.
- Conesa, M.R., Conejero, W., Vera, J., Ruiz-Sánchez, M.C. 2023. Assessment of trunk microtensiometer as a novel biosensor to continuously monitor plant water status in nectarine trees. *Frontiers in Plant Science* 14:1123045. doi:10.3389/fpls.2023.1123045.
- de la Fuente-Sáiz, D., Ortega-Farías, S., Suarez, E., Lisperguer, M.J., Tian, F., Liu, Y., et al. 2026. Water requirements and energy balance components of a drip-irrigated hazelnut (*Corylus avellana*) orchard in a Mediterranean semi-arid climate. *European Journal of Agronomy* 178:128151. doi:10.1016/j.eja.2026.128151.
- Dong, Y., Tucker, S., Singh, G., Ali, N., Yazdanpanah, N., Weide, J.V., et al. 2025. Optimizing soil moisture sensor placement through spatial variability analysis in orchards. *Smart Agricultural Technology* 12:101273. doi:10.1016/j.atech.2025.101273.
- Garofalo, S.P., Giannico, V., Costanza, L., Alhajj Ali, S., Camposeo, S., Lopriore, G., et al. 2024. Prediction of stem water potential in olive orchards using high-resolution planet satellite images and machine learning techniques. *Agronomy* 14:1. doi:10.3390/agronomy14010001.
- González-Dugo, V. 2024. Crop water stress sensing for improved management in orchards and vineyards. *Acta Horticulturae* 1395:1-6. doi:10.17660/ActaHortic.2024.1395.1.
- González-Dugo, V., Zarco-Tejada, P.J., Intrigliolo, D.S., Ramírez-Cuesta, J.M. 2021. Normalization of the crop water stress index to assess the within-field spatial variability of water stress sensitivity. *Precision Agriculture* 22:964-985. doi:10.1007/s11119-020-09768-6.
- Guo, J.S., Bush, S.E., Hultine, K.R. 2022. Temporal variation in stomatal sensitivity to vapour pressure deficit in western riparian forests. *Functional Ecology* 36:1599-1611. doi:10.1111/1365-2435.14066.
- Hoppe, H., Dietrich, P., Marzahn, P., Weiss, T., Nitzsche, C., Freiherr von Lukas, U., et al. 2024. Transferability of machine learning models for crop classification in remote sensing imagery using a new test methodology: A study on phenological, temporal, and spatial influences. *Remote Sensing* 16:1493. doi:10.3390/rs16091493.
- Jalakas, P., Takahashi, Y., Waadt, R., Schroeder, J.I., Merilo, E. 2021. Molecular mechanisms of stomatal closure in response to rising vapour pressure deficit. *New Phytologist* 232:468-475. doi:10.1111/nph.17592.
- Jenkins, M., Block, D.E. 2024. A review of methods for data-driven irrigation in modern agricultural systems. *Agronomy* 14:1355. doi:10.3390/agronomy14071355.
- Khare, D., Selzner, T., Leitner, D., Vanderborcht, J., Vereecken, H., Schnepf, A. 2022. Root system scale models significantly overestimate root water uptake at drying soil conditions. *Frontiers in Plant Science* 13:798741. doi:10.3389/fpls.2022.798741.
- Lakso, A.N., Santiago, M., Stroock, A.D. 2022. Monitoring stem water potential with an embedded microtensiometer to inform irrigation scheduling in fruit crops. *Horticulturae* 8:1207. doi:10.3390/horticulturae8121207.
- Nevo, E., Rotbart, N., Slutsky, A., Wallach, R., Gal, Y., Naor, A., et al. 2025. Using stem water potential derived from continuous reading sensors for irrigation scheduling: Nectarine as a test case. *Irrigation Science* 43:1147-1158. doi:10.1007/s00271-024-00983-x.
- Orynbaikyzy, A., Gessner, U., Conrad, C. 2022. Spatial transferability of random forest models for crop type classification using Sentinel-1 and Sentinel-2. *Remote Sensing* 14:1493. doi:10.3390/rs14061493.
- Páez-Bimos, S., Villacís, M., Morales, O., Calispa, M., Molina, A., Salgado, S., et al. 2022. Vegetation effects on soil pore structure and hydraulic properties in volcanic ash soils of the high Andes. *Hydrological Processes* 36:e14678. doi:10.1002/hyp.14678.
- Planet Labs PBC. 2026. Analysis-Ready PlanetScope. Planet Labs PBC, San Francisco, California, USA. Available at <https://docs.planet.com/data/imagery/arps/> (accessed July 2026).
- Ramírez-Cuesta, J.M., Ortuño, M.F., González-Dugo, V., Zarco-Tejada, P.J., Parra, M., Rubio-Asensio, J.S., et al. 2022. Assessment of peach trees water status and leaf gas exchange using on-the-ground versus airborne-based thermal imagery. *Agricultural Water Management* 267:107628. doi:10.1016/j.agwat.2022.107628.
- Rouault, P., Courault, D., Flamain, F., Pouget, G., Doussan, C., Lopez-Lozano, R., et al. 2024. High-resolution satellite imagery to assess orchard characteristics impacting water use. *Agricultural Water Management* 295:108763. doi:10.1016/j.agwat.2024.108763.
- Rovira, M. 2021. Advances in hazelnut (*Corylus avellana* L.) rootstocks worldwide. *Horticulturae* 7:267. doi:10.3390/horticulturae7090267.

- Sahu, H., Garg, P.K., Vijay, S., Dasgupta, A. 2026. Predictive drivers and transferability of multi-scale machine learning based crop yield prediction under drought across European and Asian climates. *Agricultural Water Management* 327:110254. doi:10.1016/j.agwat.2026.110254.
- Sandonís-Pozo, L., Llorens, J., Escolà, A., Arnó, J., Pascual, M., Martínez-Casasnovas, J.A. 2022. Satellite multispectral indices to estimate canopy parameters and within-field management zones in super-intensive almond orchards. *Precision Agriculture* 23:2040-2062. doi:10.1007/s11119-022-09956-6.
- Savchik, P., Nocco, M., Kisekka, I. 2024. Mapping almond stem water potential using machine learning and multispectral imagery. *Irrigation Science* 43:105-120. doi:10.1007/s00271-024-00932-8.
- Termin, D., Linker, R., Baram, S., Raveh, E., Ohana-Levi, N., Paz-Kagan, T. 2023. Dynamic delineation of management zones for site-specific nitrogen fertilization in a citrus orchard. *Precision Agriculture* 24:1570-1592. doi:10.1007/s11119-023-10008-w.
- Velázquez-Chávez, L.J., Daccache, A., Mohamed, A.Z., Centritto, M. 2024. Plant-based and remote sensing for water status monitoring of orchard crops: Systematic review and meta-analysis. *Agricultural Water Management* 303:109051. doi:10.1016/j.agwat.2024.109051.
- Xing, Y., Wang, X. 2024. Precise application of water and fertilizer to crops: Challenges and opportunities. *Frontiers in Plant Science* 15:1444560. doi:10.3389/fpls.2024.1444560.
- Zhu, W., Zhao, D., Di, N., Li, D., Zhou, O., Sun, Y., et al. 2023. Matching root water uptake patterns to fine root and soil water distributions. *Plant and Soil* 495:499-516. doi:10.1007/s11104-023-06349-0.