

Contents

- Table S1: Data availability during the three analysed growing seasons.
- Figure S1: Temporal distribution of data availability during the three analysed growing seasons.
- Figure S2: Distribution of 15-min mean wind speed and thresholds used to define wind-speed classes.
- 5 – Figure S3: Pairwise projections of the first four principal components of the seismic network.
- Figure S4: Distribution of selected XGBoost hyperparameter values across model configurations.
- Figure S5: Held-out test performance of the four nested predictor configurations across tree–station pairs.
- Figure S6: Absolute predictive performance, incremental seismic gains, and matched training and test sample availability.
- Figure S7: TWD variability versus incremental seismic predictive gains.

Table S1. Data availability during the analysed 2024–2026 growing seasons. Coverage is the percentage of expected 15-min intervals containing a valid observation. For a seismic station, an interval is counted as available when at least one derived seismic feature is present; the unit shown is that of the underlying ground-velocity measurement. Meteorological summaries derived from the same original tower variable are grouped together. Precipitation is reported as 15-min depth in mm; other meteorological units correspond to the source variables. The soil moisture row represents the median across the measured depths at one profile. The temporal distribution of availability is shown in Fig. S1. Source-level coverage does not imply that all predictors and the target are simultaneously available for modelling.

Data source and variable	Unit	2024 (%)	2025 (%)	2026 (%)
<i>Seismic stations</i>				
Station C66	ms ⁻¹	100.0	100.0	100.0
Station C67	ms ⁻¹	100.0	98.2	72.9
Station C68	ms ⁻¹	100.0	100.0	99.2
Station C69	ms ⁻¹	100.0	100.0	92.4
Station C6A	ms ⁻¹	100.0	100.0	56.8
<i>Tree water deficit observations</i>				
Tree Fm1 (European beech)	µm	0.0	96.1	41.4
Tree Fm2 (European beech)	µm	100.0	92.8	93.2
Tree Fm4 (European beech)	µm	100.0	96.1	93.0
Tree Fm5 (European beech)	µm	100.0	96.1	99.6
Tree Pm1 (Douglas fir)	µm	99.9	83.2	99.6
Tree Pm2 (Douglas fir)	µm	100.0	95.3	99.6
Tree Pm4 (Douglas fir)	µm	100.0	96.1	99.6
Tree Pm5 (Douglas fir)	µm	100.0	96.1	99.6
<i>Soil moisture</i>				
Median across 0.3, 0.5, 0.7, and 0.9 m depths	%	98.7	95.2	99.1
<i>Meteorological tower</i>				
Air temperature, 2 m	°C	99.9	96.1	99.6
Air temperature, 46 m	°C	99.8	96.8	99.6
Atmospheric pressure, 2 m	kPa	100.0	96.1	99.6
Precipitation depth, 47 m	mm	99.8	96.8	99.6
Relative humidity, 2 m	%	100.0	96.8	99.6
Relative humidity, 46 m	%	99.9	96.8	99.6
Wind direction, 47 m	°	99.9	96.8	99.6
Wind speed, 47 m	ms ⁻¹	99.9	96.8	99.6
Maximum wind speed, 47 m	ms ⁻¹	99.9	96.8	99.6

10 **S1 Data availability**

Data coverage was assessed relative to the complete expected 15-min temporal grid of each analysed growing season: 5 July–30 September 2024, 1 May–30 September 2025, and 1 May–30 August 2026. Table S1 summarizes coverage over each season, while Fig. S1 shows the temporal distribution of available observations and data gaps. These source-level coverage estimates differ from the numbers of matched samples retained for modelling, which are shown in Fig. S6.

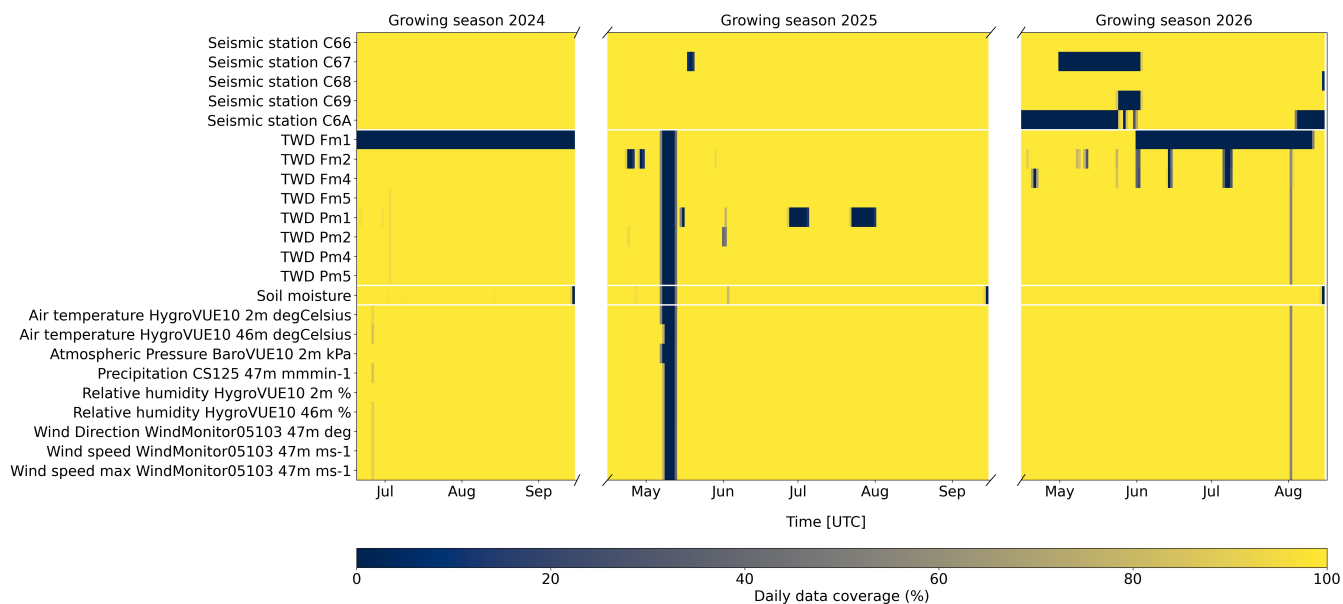


Figure S1. Temporal data availability during the analysed 2024–2026 growing seasons. Rows represent the five seismic stations, eight dendrometer trees, the median soil moisture series, and meteorological variables. Colour indicates the percentage of valid 15-min intervals within each day, from no available data (dark blue) to complete daily coverage (yellow). White horizontal lines separate data sources. Breaks in the time axis omit the intervals between growing seasons. Corresponding coverage percentages aggregated over each season are reported in Table S1.

15 S2 Wind-speed distribution and class thresholds

Wind-conditioned analyses use three classes based on 15-min mean wind speed: low ($< 2 \text{ m s}^{-1}$), moderate ($2\text{--}5 \text{ m s}^{-1}$, including both thresholds), and high ($> 5 \text{ m s}^{-1}$). Figure S2 shows these thresholds relative to the wind-speed distribution across the analysed growing seasons.

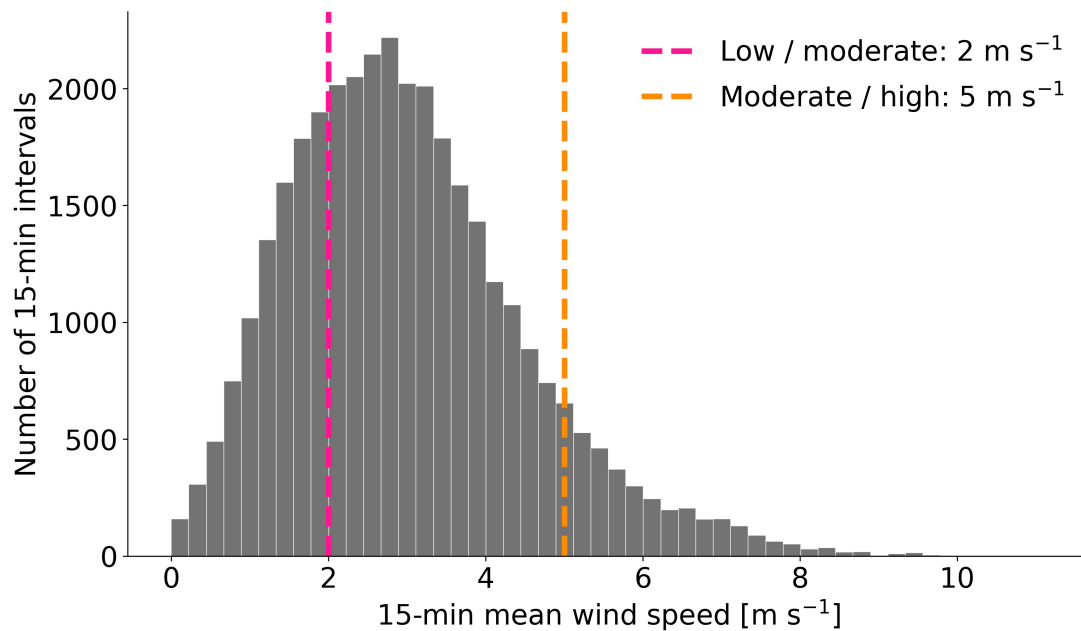


Figure S2. Distribution of 15-min mean wind speed measured at 47 m on the meteorological tower during the analysed 2024–2026 growing seasons. Bars show the number of valid 15-min intervals. Dashed lines mark the thresholds separating low ($< 2 \text{ m s}^{-1}$), moderate ($2\text{--}5 \text{ m s}^{-1}$), and high ($> 5 \text{ m s}^{-1}$) wind speeds. Observations outside the analysed growing seasons are excluded.

S3 Additional network state-space projections

- 20 To examine structure beyond the leading PC1–PC2 representation, Fig. S3 shows all pairwise projections of the first four principal components. The network mean was subtracted at each time step before median imputation, standardization, and principal component analysis, following the workflow described in the main text. The first four components explain 99.6% of the variance.

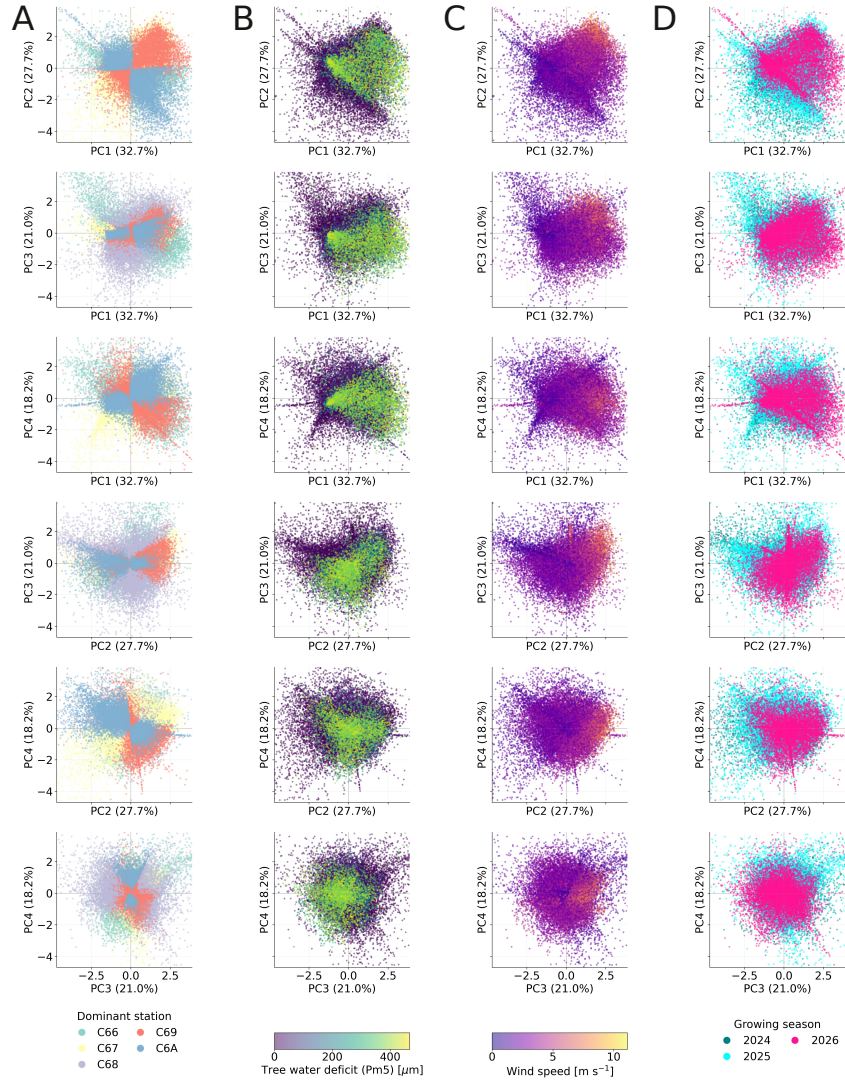


Figure S3. Pairwise projections of the first four principal components derived from the five-station seismic spectral-centroid frequencies. Rows show, from top to bottom, PC1–PC2, PC1–PC3, PC1–PC4, PC2–PC3, PC2–PC4, and PC3–PC4. PC1, PC2, PC3, and PC4 explain 32.7%, 27.7%, 21.0%, and 18.2% of the variance, respectively. Columns show observations coloured by (A) the dominant contributing seismic station, (B) TWD of Douglas fir Pm5, (C) wind speed, and (D) growing season. The panels display overlapping distributions rather than discrete TWD-defined groups.

S4 Hyperparameter selection

- 25 Hyperparameters were selected independently for each fitted model using group-based cross-validation on its training blocks, as described in the Methods. Figure S4 summarizes the selected values across tree–station combinations for each model configuration. It describes selection frequencies within the evaluated grid, not model performance on the held-out test blocks.

Model	n_estimators		max_depth			learning_rate		subsample		colsample_bytree	
	100	200	3	4	6	0.03	0.05	0.8	1	0.8	1
METEO	70%	30%	72%	25%	2%	70%	30%	60%	40%	48%	52%
METEO + SOIL	70%	30%	80%	15%	5%	70%	30%	55%	45%	57%	42%
METEO + SEISMO	68%	32%	75%	22%	2%	75%	25%	62%	38%	35%	65%
METEO + SOIL + SEISMO	45%	55%	70%	12%	18%	70%	30%	50%	50%	60%	40%
METEO + SOIL + RESID	98%	2%	90%	10%	0%	98%	2%	38%	62%	82%	18%
SEISMO	98%	2%	92%	8%	0%	100%	0%	55%	45%	75%	25%
Selected value											

Figure S4. Distribution of selected XGBoost hyperparameter values across fitted tree–station models. Rows correspond, from top to bottom, to METEO, METEO + SOIL, METEO + SEISMO, METEO + SOIL + SEISMO, METEO + RESID, and SEISMO. The row labelled METEO + RESID reports the hyperparameters of the fitted seismic residual model in the residual-correction workflow. The METEO + SOIL + RESID row reports the hyperparameters of the seismic residual model correcting the METEO + SOIL baseline. Columns show the number of estimators, maximum tree depth, learning rate, row subsampling fraction, and feature subsampling fraction. Each cell gives the percentage of fits for which the corresponding value was selected within that model configuration. Percentages are rounded to whole numbers and may therefore not sum to exactly 100%. Hyperparameter selection uses training blocks only.

S5 Absolute predictive performance across tree–station pairs

Figure S5 provides the absolute held-out test R^2 values for the four nested predictor configurations across all 40 tree–station pairs. These values complement the incremental gains shown in Fig. 5 of the main text: the performance of a model containing seismic predictors should be considered alongside that of its corresponding baseline model.

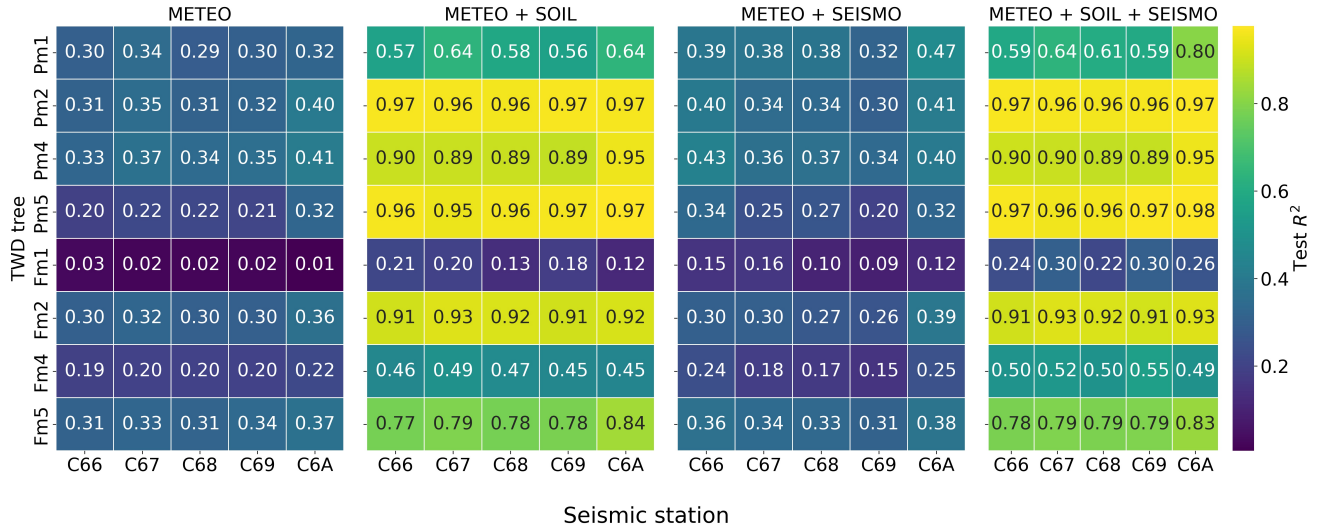


Figure S5. Held-out test R^2 across 40 tree–station combinations for four predictor configurations, shown from left to right: METEO, METEO + SOIL, METEO + SEISMO, and METEO + SOIL + SEISMO. Rows identify trees (Pm: Douglas fir; Fm: European beech), and columns identify seismic stations. Metrics are calculated on held-out observations pooled across the analysed 2024–2026 growing seasons. Each nested comparison uses matched observations for the corresponding tree–station pair. Values are rounded to two decimals, and all four panels use the same colour scale. Models without seismic predictors are evaluated within each pair-specific matched dataset; their results may therefore differ between station columns for the same tree. Incremental seismic gains in R^2 and relative RMSE reduction are shown in Fig. 5 of the main text.

S6 Predictive performance and matched sample availability

Figure S6 places absolute predictive performance and incremental seismic gains alongside the numbers of training and test observations retained for each tree–station pair. Unlike the source-level coverage in Table S1 and Fig. S1, these counts describe 15-min samples with jointly available targets and predictors retained for the METEO + SOIL + SEISMO model. The figure provides descriptive context for differences between pairs; it does not establish a causal effect of sample availability on predictive performance.

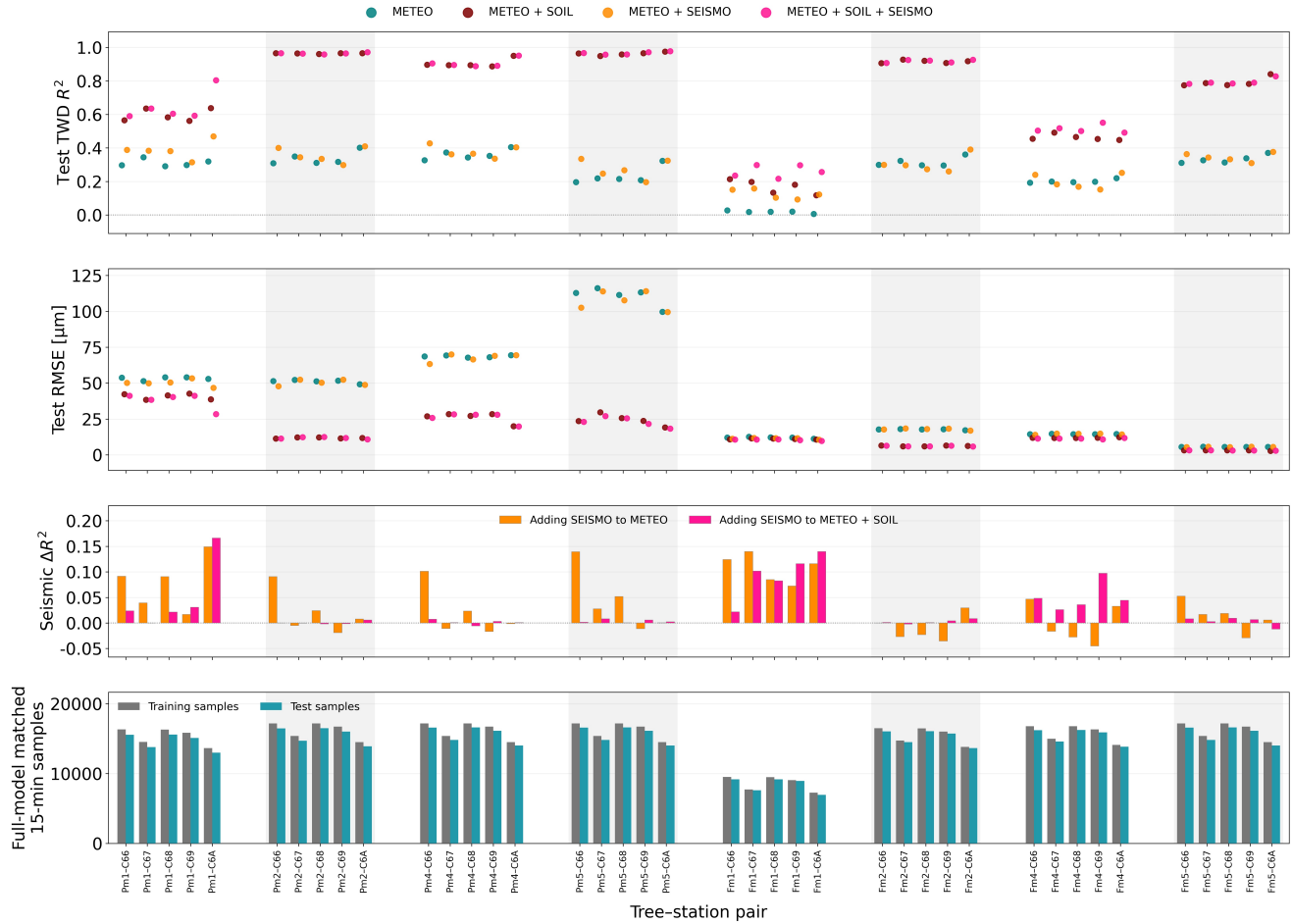


Figure S6. Held-out predictive performance, incremental seismic gains, and matched sample availability across 40 tree–station combinations. The top two panels show test R^2 and RMSE (in μm), respectively, for METEO (teal), METEO + SOIL (maroon), METEO + SEISMO (orange), and METEO + SOIL + SEISMO (pink). Metrics are pooled across the analysed 2024–2026 growing seasons. Points are displaced slightly along the horizontal axis to distinguish models within each pair. The third panel shows the change in held-out R^2 when seismic predictors are added to METEO (orange) or METEO + SOIL (pink). Positive values indicate improved predictive performance; negative values indicate deterioration. The bottom panel shows the numbers of matched 15-min training samples (grey) and test samples (blue-green) retained for the METEO + SOIL + SEISMO model. These counts describe usable modelling observations, not total dendrometer coverage or independent replicates. Alternating background shading distinguishes trees.

S7 TWD variability and incremental seismic predictive gains

40 TWD variability was quantified separately for each tree–station pair as the difference between the 90th and 10th percentiles of observed TWD on its matched held-out test observations. This robust range describes the spread of the target values represented in the evaluation dataset. It measures variability rather than the level or severity of water deficit. Figure S7 compares this range with the incremental predictive gains from adding seismic features. Tree–station pairs share trees and stations, so the observations are not treated as independent replicates for pooled significance testing.

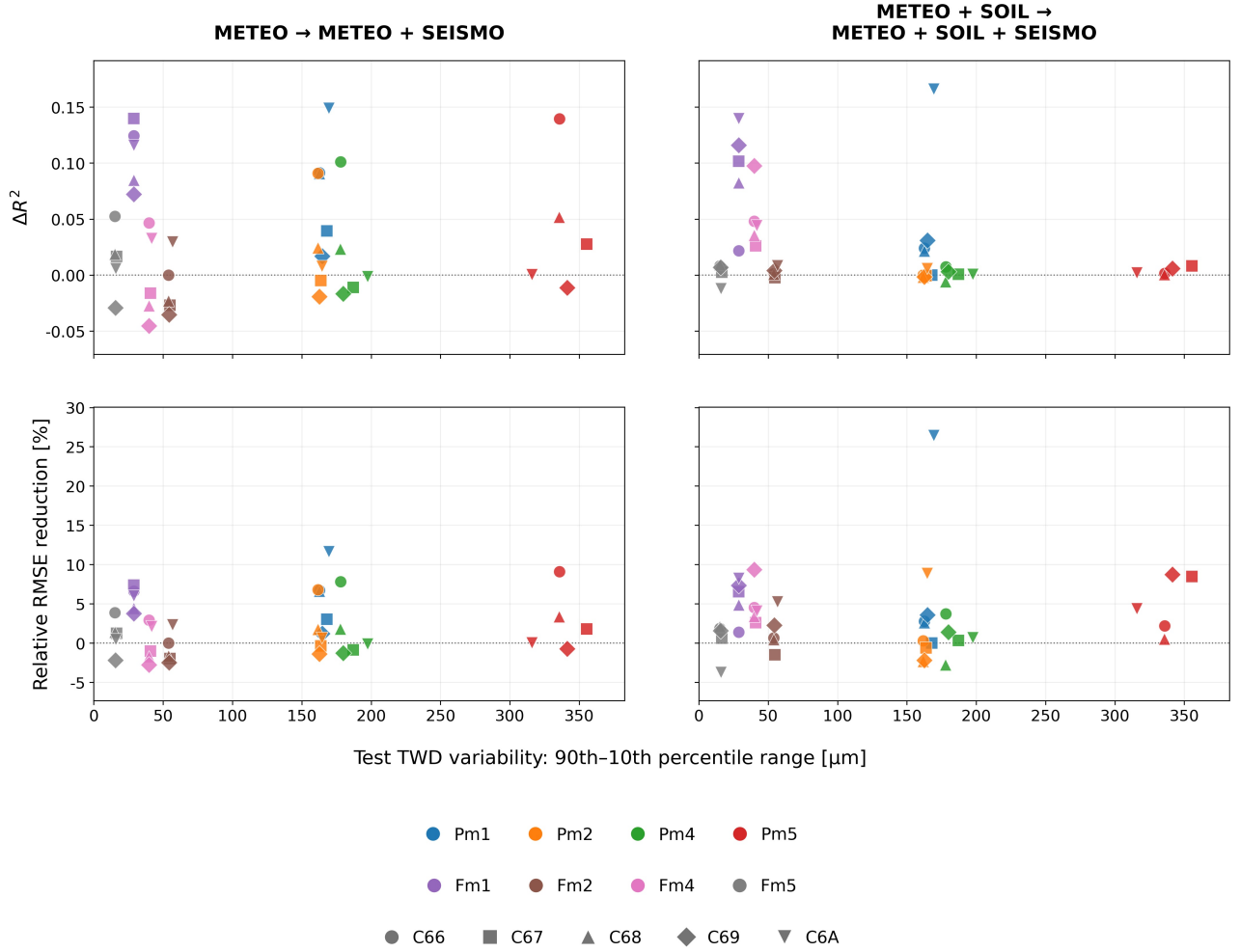


Figure S7. Observed TWD variability versus the added predictive value of seismic features across 40 tree–station combinations. The horizontal axis shows the difference between the 90th and 10th percentiles of observed TWD (in μm), calculated on matched held-out test observations pooled across the analysed growing seasons. The left column shows the comparison METEO → METEO + SEISMO, and the right column shows METEO + SOIL → METEO + SOIL + SEISMO. The top row shows the change in held-out R^2 ; the bottom row shows relative RMSE reduction, calculated as $100 \times (\text{RMSE}_{\text{baseline}} - \text{RMSE}_{\text{combined}}) / \text{RMSE}_{\text{baseline}}$. Positive values indicate improved predictive performance; negative values indicate deterioration. Each point represents one tree–station pair. Colours identify individual trees, and marker shapes identify seismic stations. The TWD range can differ between stations for the same tree because their matched evaluation periods differ. Dotted horizontal lines mark zero gain. The plots are descriptive: TWD variability is not a measure of drought severity, and points sharing trees or stations are not independent replicates.