Robustness of high-throughput prediction of leaf ecophysiological traits using near infra-red spectroscopy and poro-fluorometry.

Eva Coindre[1,2], Romain Boulord[2], Laurine Chir[2], Virgilio Freitas[1,2], Maxime Ryckewaert[3], Thomas Laisné[2], Virginie Bouckenooghe[4,5], Maëlle Lis[2], Llorenç Cabrera-Bosquet[2], Agnès Doligez[1,5], Thierry Simonneau[2], Benoît Pallas[2], Aude Coupel-Ledru[2]*, Vincent Segura[1,5]*.

[1] UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France;
[2]LEPSE, Univ Montpellier, INRAE, Institut Agro, Montpellier, France;
[3]Inria, LIRMM, Univ Montpellier, CNRS, Montpellier, France;
[4]Institut français de la vigne et du vin, Pôle National Matériel Végétal, Le Grau du Roi 30240, France
[5]Geno-Vigne®, IFV-INRAE-Institut Agro, F-34398, Montpellier, France;
(*Authors for correspondence: email aude.coupel-ledru@inrae.fr and vincent.segura@inrae.fr)

Abstract

Background: Water resource is a major limiting factor impacted by climate change that threatens crop production and quality. Understanding the ecophysiological mechanisms involved in the response to water deficit is crucial to select new varieties more drought tolerant. A major bottleneck hampering such advances is the lack of methods for measuring ecophysiological traits on large populations of individuals. We investigated the relevance of spectroscopy and poro-fluorometry with related high-throughput measurement devices to predict leaf morphological and ecophysiological traits using partial least square regression (PLSR) models. This work relies on a grapevine diversity panel grown in pots under contrasted conditions, outdoors under well watered conditions and in a greenhouse with three different soil water treatments. We took advantage of these experimental designs to specifically assess the robustness of predictive models.

Results: Some complementarity between measuring devices were found, with spectrometers being able to predict leaf mass per area, water content and water quantity (R^2>0.58), while the poro-fluorometer could predict net CO_2 assimilation (R^2>0.72), regardless of the water treatment. The prediction of leaf mass per area appeared to be quite robust between outdoors and greenhouse experiments. The prediction of water use efficiency was highly dependent on the water treatment, with much better predictions under moderate (R^2>0.73) than severe water deficit. We then used the high-throughput devices alone to measure the whole grapevine panel, by applying calibration models to predict ecophysiological traits and estimated their broad sense heritability. Leaf mass per area was also directly determined on the entire diversity panel and its

heritability was similar whether calculated on observed or predicted values. The highest computed heritabilities for these traits reached values close to 0.5.

Conclusions: This study showed that spectrometers and poro-fluorometer could be reliable nondestructive tools for the high-throughput phenotyping of ecophysiological traits on thousands of plants, paving the way for studying the genetic determinism of such traits.

Keywords : High-throughput phenotyping, near-infrared spectroscopy (NIRS), fluorescence, Partial least square regression (PLSR), leaf traits, diversity panel, soil water deficit, grapevine, broad-sense heritability

Introduction

Over the last decades, climate change has become a global concern. The Mediterranean region is particularly concerned, where prediction expects significant temperature rises and changes in annual rainfall patterns [1,2]. In order to maintain crops in this context, there is a need to explore new combinations of genetic and management choices depending on environmental conditions. Identifying the best performing solutions further requires high throughput characterization of traits related to plant functioning and their responses to the environment.

Gas exchange, leaf structure and specific leaf biochemical components are among the traits most impacted by abiotic constraints such as soil water availability. The regulation of gas exchange at the leaf level by stomatal closure/opening plays a key role in the whole plant response to water stress. Indeed, a soil water deficit leads to a drop in water potential all along the plant that could damage its water transfer capacity and result in organ or plant death [3]. To prevent such damage, plants can quickly close their stomata, reducing transpiration (E) [4] and helping to maintain a high leaf water content (WCf) [5]. However, this water saving strategy has a heavy cost for the plant as it also decreases net photosynthesis (A_n) by decreasing CO_2 uptake [3], and in turn biomass production. A trade-off between reducing E and maintaining A_n under water stress can be sought by considering the water use efficiency, *i.e.* the biomass produced per unit of water consumed. At the leaf level, it is measured with the instantaneous water use efficiency (WUE_{inst}), calculated as the ratio between A_n and E, while the intrinsic water use efficiency (WUE_{intr}),

63 calculated as the ratio between A_n and the stomatal conductance for water vapor (g_{sw}), provides
64 an alternative estimate, less dependent on air vapor pressure deficit and closer to the genetic
65 determinisms than WUE_{inst} [6].

66 In grapevine, the species of interest in the present work, a large genotypic variability has already
67 been observed for A_n, g_{sw}, WUE_{intr}, leaf water potential (Ψ) [7–10]. However, this variability was
68 observed for a limited number of cultivars and/or in a limited number of growing conditions, usually
69 well-watered conditions. High-throughput assessment of these ecophysiological traits is limited
70 by the time required for measurements using leaf gas analyzers or pressure chambers. Studying
71 large populations with a wide range of genetic diversity under various conditions is crucial to
72 identify the best performing cultivars. It can also enable the identification of genetic regions
73 associated with these traits and their responses to contrasted environmental conditions [10,11].
74 A main challenge is thus to investigate the use of alternative high-throughput methods as proxies
75 to accurately predict traits of interest on large panels of genotypes under various water availability
76 treatments and under different growing conditions [12].

77 Fluorometer devices, measuring the quantum efficiency of photosystem II, electron transport rate,
78 and photochemical and non-photochemical quenching coefficients, have been proposed as a
79 low-cost and fast method to estimate A_n [13,14]. Chlorophyll fluorescence measurements were
80 also reported to evaluate the variability of photosynthetic performance among a population of
81 hybrids in barley [15]. Moreover, previous studies have calculated a semi-empirical index that
82 combines fluorescence parameters and leaf temperature, to estimate net photosynthesis in fruit
83 trees [16]. This indicator was then used to evaluate the photosynthesis activity of a core-collection
84 of apples in response to water deficit [17]. Near-infrared spectroscopy (NIRS) is another high-
85 throughput method commonly used to predict plant tissue composition [18–20]. Dedicated
86 chemometrics methods, such as Partial Least Square Regression (PLSR) [21], are used to
87 establish relationships between spectral absorbance or reflectance and reference measurements

88 obtained mostly by low-throughput methods. NIRS has been reported in plants to predict
89 numerous leaf compositional traits, including nitrogen [22] and carbon content, lignin, cellulose
90 [19], protein, starch, total non-structural carbohydrates [23] and sugars content [24,25]. NIRS has
91 also been implemented to predict various morphological and ecophysiological traits on multiple
92 species and on large populations of the same species. For instance, leaf mass per area (LMA),
93 which is widely used in plant ecology as an indicator of plant functioning, has been predicted by
94 NIRS on diverse species [24,26]. NIRS predictive models were also reported for LMA and A_n on
95 wheat and triticale germplasms [27] as well as on hybrids of maize [28]. LMA, A_n, WUE_{intr}, WUE_{inst},
96 Ψ and leaf water content were together predicted by NIRS on a diversity panel of maize [22] and
97 Ψ has also been predicted by NIRS on several cultivars of grapevine [29,30]. The relevance of
98 NIRS to predict water related traits was observed to be associated with light absorption by water
99 at wavelength bands centered around 1400 and 1900 nm [31,32]. However, the presence of water
100 can also reduce leaf reflectance, hindering the detection of other components. To minimize the
101 impact of water on spectra and consequently on the prediction accuracy of other traits, spectra
102 can be collected on dried leaves [33]. Alternatively, the water absorption effect can be reduced
103 using water-removed spectra [34] or by using an external parameter orthogonalisation [35].
104 Finally, not all spectrometers measure the same wavelength range. Therefore, the type of
105 spectrometer must be selected according to its application [36,37].

106 When studying ecophysiological traits, which vary according to experimental conditions, the
107 robustness of PLSR model predictions must be examined. On grapevine, predictive models with
108 NIRS data have been built on few cultivars subjected to different soil water treatments [29,30,38].
109 PLSR models were calibrated and validated with datasets containing samples from all water
110 treatments [29,38], without assessing the ability of the model calibrated on a given water
111 treatment to predict the values of plants subjected to another soil water treatment. If PLSR models
112 are robust across different aerial environmental conditions and soil water availabilities,

phenotyping time could be reduced in new experiments. Another issue with the PLSR model is the risk of overfitting. To prevent this risk, the calibration model is usually built with a cross-validation procedure. However, even in a cross-validation setting, the predictive ability of the model could still depend on the calibration and validation sets that could be randomly or non-randomly chosen [39].

In this study, we evaluated the ability and complementarity of NIRS and poro-fluorometry data to predict leaf traits that are difficult to access at high-throughput. These two innovative approaches were applied on a grapevine diversity panel [40], grown in pots both outdoors under non-limiting water availability and in a greenhouse with varying intensities of soil water deficit. The robustness of PLSR models was studied across experiments and water treatments. Heritabilities of traits either directly measured or estimated from high-throughput data were compared to investigate the ability of high-throughput phenotyping to assess the genotypic variability.

Material and methods

Plant material and experimental design

We used a subset of 241 cultivars from a panel of 279, previously designed to represent the worldwide genetic diversity of *Vitis vinifera* [40]. We added 5 cultivars that are in the process of being registered in the French catalog, resulting in 246 studied cultivars. Twenty own-rooted plants were grown for each cultivar with ferti-irrigation from 2018 at the experimental vineyard of l'Institut Agro Montpellier (France) in individual 3-liter pots containing a 30:70 (v/v) mixture of loamy soil and organic compost [41]. In January 2020, the plants were transferred into larger, individual 9-liter pots filled with the same substrate.

A first phenotyping season was conducted in July 2021, on the potted plants grown in the experimental vineyard under non-limiting, drip-irrigated conditions ("Outdoor experiment"). Leaf

phenotyping was performed during seven mornings on two to three plants per cultivar. Six plants out of the 20 available per cultivar were next retained to be part of a greenhouse experiment in 2022, where different watering treatments were set up ("Greenhouse experiment"). This second experiment was conducted in the PhenoArch phenotyping platform [42] hosted at the M3P (Montpellier Plant Phenotyping Platforms, https://eng-lepse.montpellier.hub.inrae.fr/platforms-m3p/montpellier-plant-phenotyping-platforms-m3p), from May to August 2022. The whole panel (246 cultivars) was subjected to three different water treatments: Well-Watered (WW), moderate Water Deficit (WD1) and severe Water Deficit (WD2) maintained at different soil water content in the pots, with two plants for each cultivar and each water treatment (1476 plants in total). These treatments correspond to soil water potential equal to -0.03 MPa (well watered conditions, WW), -0.11 Mpa (moderate water deficit, WD1) and -0.43 Mpa (severe water deficit, WD2) (protocol for management of water treatment as in Coupel-Ledru et al., 2014 [10]). For each cultivar, the six plants were grouped into two triplets (one triplet including one WW, one WD1 and one WD2 plant), and cultivar triplets were randomized within the phenotyping platform. Leaf phenotyping was conducted over 24 mornings in June and July 2022, once the water content in the pots had stabilized at their target values. Environmental variables were monitored using 10 probes, strategically distributed throughout the greenhouse.

Leaf phenotyping

For both experiments (outdoors and greenhouse), measurements were performed between 9:00 and 13:00 (UTC+2), maximizing environmental and traits variations to support the statistical calibration process. Leaves were chosen as fully expanded and well exposed to sunlight. They were usually located between the tenth and the fifteenth phytomer from the apex. Two kinds of measurements were performed: high-throughput phenotyping on all six plants of the 246 cultivars and low-throughput phenotyping, on a subset of randomly selected plants. Multiple devices were

160 used successively on the same leaf *in situ*, in a short time interval, for both high- and low-
161 throughput phenotyping. The first device used was a hand-held poro-fluorometer (Li600, Portable
162 Photosynthesis System LI-600, Licor, Lincoln, United States). It was used to measure the
163 transpiration rate, E (mmol m^{-2} s^{-1}), the stomatal conductance g_{sw} (mol m^{-2} s^{-1}), chlorophyll
164 fluorescence parameters and micro-climate variables i.e. leaf vapor pressure deficit, atmospheric
165 pressure and ambient light (Supplementary Table 1). Two measurements were taken at two
166 different points on the same leaf and were averaged to obtain a single value per leaf. The second
167 and third devices were micro-portable near infrared spectrometers: MicroNIR (MicroNIR™ Onsite-
168 W, VIAVI, Chandler, AZ, United States) and NeoSpectra (NeoSpectra™ Scanner, Si-Ware,
169 Miami, FL, United States). The MicroNIR covers a spectral range from 908 nm to 1676 nm with
170 125 wavelengths, resulting in a spectral sampling interval of 6.2 nm. The NeoSpectra covers a
171 spectral range from 1350 nm to 2500 nm with 257 wavelengths and a spectral sampling interval
172 of 16 nm. The wavelength interval between 1350 nm and 1676 nm was common between both
173 devices. For both devices, two spectra were collected per leaf on the lamina while avoiding the
174 principal veins. These two spectra were respectively collected on each side of the leaf (abaxial
175 and adaxial faces) and were then averaged to obtain a single spectrum per leaf. After these
176 measurements, a 3 cm diameter disk was sampled from the leaf. In the greenhouse experiment,
177 this disk was immediately weighed after sampling to obtain its fresh weight (FW, mg) with 0.1 mg
178 accuracy (Sartorius balance, QUINTIX224-1S, Göttingen, Germany). For both experiments, the
179 disk was oven-dried for one week at 60°C before being weighed to get its dry weight (DW, mg).
180 Leaf Mass per Area (LMA, mg cm^{-2}) was calculated as the ratio between DW and disk area (7.07
181 cm^2). For disks of leaves collected in the greenhouse experiment only, leaf water content (WC_f,
182 mg H_2O mg^{-1} fresh weight) was also calculated, as $\frac{FW - DW}{FW}$ and the water quantity (WQ, mg H_2O
183 per leaf disk) as $FW - DW$. The measurement of LMA, WCf, and WQ is traditionally considered a
184 low-throughput method due to the significant time and sample storage requirements. However,

185 in this study, these traits were measured on all plants in the greenhouse experiment. The high-
186 throughput leaf phenotyping was performed during 24 mornings on the greenhouse experiment.
187 After weighing, dried disks were stored at 20°C. They were later used again to collect NIRS
188 spectra using a laboratory spectrometer ASD (ASD LabSpec® 4, Analytical Spectral Devices,
189 Boulder, USA). The temperature was maintained at 20°C in the lab during measurements. Four
190 spectra were taken using a reflectance probe (4 mm diameter) at different points of the dried disk
191 before being averaged. The ASD covers a spectral range from 350 nm to 2500 nm with 2150
192 wavelengths and a sampling interval of 1 nm.
193 The low-throughput measurements presented in this paper were performed only in the
194 greenhouse experiment on a subset of 120 plants (34 cultivars) subjected to three water
195 treatments. They were performed during 4 mornings (June 14, 2022, July 6, 2022, July 21, 2022
196 and July 26, 2022). The measurements started as described above for the high-throughput
197 devices (i.e. Li600, MicroNir, NeoSpectra), and were followed by a measure with an InfraRed
198 Gas-exchange Analyzer (IRGA) equipped with a light-emitting diode light source (Li6800, Portable
199 Photosynthesis System, Licor, Lincoln, United States). This device allowed measuring net CO_2
200 assimilation (A_n, µmol m^{-2} s^{-1}), transpiration rate (E, mmol m^{-2} s^{-1}) and stomatal conductance (g_{sw},
201 mol m^{-2} s^{-1}). Intrinsic Water Use Efficiency (WUE_{intr}, µmol CO_2 mol^{-1} H_2O) and instantaneous
202 Water Use Efficiency (WUE_{inst}, µmol CO_2 $mmol^{-1}$ H_2O) were further calculated as the ratio
203 between A_n and g_{sw} or A_n and E, respectively. The conditions in the leaf chamber during the
204 measurement were set to 400 µmol mol^{-1} CO_2, with actinic light adjusted to a value corresponding
205 to the ambient Photosynthetically Photon Flux Density (PPFD) measured simultaneously by the
206 Li600 device for each measured leaf. PPFD values thus ranged from 110 to 1150 µmol m^{-2} s^{-1}.
207 The leaf water potential (Ψ, MPa) was next measured using a pressure chamber (Scholander
208 pressure bomb, SoilMoisture Equipment Corp., Goleta, United States). Finally, leaf disks were
209 sampled and oven-dried to access LMA and collect spectra with the ASD device.

Preprocessing of spectral and fluorescence data

All data analyses were handled in R v4.3.0 [43]. Spectra from the ASD laboratory spectrometer were initially processed using the rmgap function (R/rchemo package) [44] to remove vertical gaps caused by differences between detectors. Wavelengths between 350 and 400 nm were removed and excluded from spectral analyses due to their noisy signals [45]. The spectra were then smoothed using R/loess, with the neighborhood size controlled by an α value of 0.03. In order to express spectra from all devices in reflectance and wavelength (nm), the MicroNIR absorbance (A) was converted into reflectance (R), as $R = \frac{1}{10^A}$, and the NeoSpectra wavenumber (v, cm^{-1}) was converted into wavelength (λ, nm), as $\lambda = \frac{10^7}{v}$. As the MicroNIR and NeoSpectra spectrometers have complementary ranges, with an overlap from 1350 nm to 1676 nm, a new, additional spectrum called MicroNIR_NeoSpectra was created for each leaf. Whole MicroNIR spectra from 950 nm to 1676 nm were retained due to their higher resolution than NeoSpectra spectra and combined with the 1676 nm - 2500 nm ranges measured with the NeoSpectra. The resulting gap at the junction of the two spectra was removed using R/rmgap from R/rchemo. All spectra (ASD, MicroNIR, NeoSpectra and MicroNIR_NeoSpectra) were then preprocessed using Standard Normal Variate (SNV) [46]. This was followed by a first-order derivative using the Savitzky-Golay filter (R/sgolayfilt from R/signal package, [47,48]). The polynomial order for the first-order derivative was set to 2, and the window sizes were set respectively to 11, 15, 15 and 41 for MicroNIR, NeoSpectra, MicroNIR_NeoSpectra and ASD, to account for their differences in resolution.

Twenty-one environmental, porometry and fluorescence measured variables were selected from the Li600 device raw measurements for further use as explicative variables in PLSR models (Supplementary Table 1).

Data modeling and cross-validation

Calibration models were constructed in order to relate high-throughput data from the spectrometers and the poro-fluorometer Li600 to the traits of interest measured with conventional, low-throughput devices or methods (IRGA, pressure chamber, disks weighing) using Partial Least Square Regression (PLSR), as implemented in R/pls with the “kernelpls” method [49]. Calibration was performed for the following traits: LMA, WC_f, WQ, A_n, Ψ, WUE_{intr} and WUE_{inst}, using data from the greenhouse experiment. For the outdoors experiment, only LMA was considered as the other traits were not measured. The number of leaves used to calibrate the model varied from 72 to 1437, depending on available data for the measured trait and the cross-validation settings (Table 1). Cross-validation was performed using either a 4-fold, 5-fold, or leave-one-out (LOO) method [50]. Four- and 5-fold cross-validations were repeated 10 times (Table 1).

PLSR models were built with a maximum number of latent variables set at 20 for spectrometers and 10 for the Li600. The optimal number of latent variables was determined by minimizing the average Root Mean Squared Error (RMSE) over the 10 iterations of the 4- and 5-fold cross-validations or directly the RMSE for the LOO cross-validation. Model quality was evaluated either within the cross-validation or using validation (testing) sets (see section below for details on the different calibration and validation sets tested) using both the R^2 computed as the squared correlation between observed and predicted values, and the RMSE.

Table 1 Strategies to evaluate the robustness and the prediction quality of PLSR models.

				Calibration							Validation				
				Nb of leaves					Cross-validation settings		Nb of leaves				Settings
	Dataset	LT or HT	Traits	WW	WD1	WD2	Outdoors	Total	Nb fold	Nb iter.	WW	WD1	WD2	Outdoors	Nb iter.
Devices complementarity	**A**	HT	LMA, WC_f, WQ	483	481	473	-	1437	5-fold	10	-	-	-	-	-
	B	LT	LMA, A_n, Ψ, WUE_{intr}, WUE_{inst}	37	36	36	-	109	5-fold	10	-	-	-	-	-
Robustness between experiments	**C**	HT	LMA	-	-	-	370	370	4-fold	10	100	100	100	100	15
	D	HT	LMA	124	123	123	-	370	4-fold	10	100	100	100	100	15
	E	HT	LMA	93	92	93	92	370	4-fold	10	100	100	100	100	15
Robustness between water treatments	**F**	HT	LMA, WC_f, WQ	370	-	-	-	370	4-fold	10	100	100	100	-	15
	G	HT	LMA, WC_f, WQ	-	370	-	-	370	4-fold	10	100	100	100	-	15
	H	HT	LMA, WC_f, WQ	-	-	370	-	370	4-fold	10	100	100	100	-	15
	I	HT	LMA, WC_f, WQ	124	123	123	-	370	4-fold	10	100	100	100	-	15
	J	LT	A_n, Ψ, WUE_{intr}, WUE_{inst}	24	24	24	-	72	LOO	10	10	10	10	-	100
Predictive models	**K**	LT	LMA, A_n, WUE_{intr}, WUE_{inst}	34	34	34	-	102	4-fold	1	-	-	-	-	1

Leaf mass per area (LMA), Water content (WC_f), Water quantity (WQ), Net CO_2 assimilation (A_n), Leaf water potential (Ψ), Intrinsic water use efficiency (WUE_{intr}), Instantaneous water use efficiency (WUE_{inst}), High-throughput phenotyping (HT), Low-throughput phenotyping (LT), Number of iterations (Nb Iter.), Leave-one out (LOO), Well-watered treatment (WW), Moderate water deficit treatment (WD1), Severe water deficit treatment (WD2).

Devices complementarity

The first analyses aimed to compare the predictive performance of PLSR models across the different devices for each trait, using only data from the greenhouse experiment. The calibration set included the same number of leaves for each water treatment, consisting respectively of approximately 480 leaves (dataset A in Table 1) or 36 leaves (dataset B in Table 1) for traits measured at high-throughput (LMA, WC_f and WQ) or low-throughput (LMA, A_n, Ψ, WUE_{intr} and WUE_{inst}), respectively. A 5-fold cross-validation was conducted based on cultivars rather than individual leaves to avoid the confounding effect of cultivar prediction on the trait prediction. PLSR models were compared based on their cross-validated R^2_{cv}, cross-validated normalized root mean square error $NRMSE_{cv}= \frac{RMSE_{cv}}{\bar{x}}$ (with $\bar{x}$ the mean of the trait) and using a Student's t test, as implemented in R/rstatix.

Robustness between experiments

LMA, was the trait used to assess the robustness of the PLSR models between experiments (oudoors or in the greenhouse). Three calibration sets were used; each with a fixed size of 370 leaves (Table 1). The first calibration set (dataset C) included only outdoors leaves, the second one (dataset D) included only greenhouse data, with approximately 123 leaves from each water treatment and the last one (dataset E) included data from both outdoors and greenhouse experiments.

The predictive ability of the PLSR models was assessed using validation sets, each including 100 leaves, selected to represent experiments and water treatments. The composition of the validation set was determined through random sampling, which accounted for the proportion of plants phenotyped each day in order to be representative of the entire phenotyping series. This sampling was repeated 15 times, resulting in 15 distinct calibration and validation sets. PLSR models were

279 compared based on their R^2 and RMSE and a Tukey HSD post-hoc test, as implemented in
280 R/stats.

281 Robustness between water treatments

282 Data from the greenhouse experiment were used to assess the robustness of the PLSR models
283 between the different water treatments. For LMA, WC_f and WQ, which were phenotyped at high-
284 throughput, 4 calibration sets were tested, each including a fixed number of 370 leaves, randomly
285 selected independently of cultivar. A separate calibration set was created for each water treatment
286 (datasets F to H in Table 1) and an additional calibration set (dataset I) that included the three
287 water treatments in equal proportions of 123 leaves was also considered.
288 The predictive ability of the PLSR models was assessed using validation sets, each including 100
289 leaves per water treatment and accounting for the proportion of plants phenotyped each day in
290 order to be representative of the entire phenotyping series. This sampling was repeated 15 times,
291 resulting in 15 distinct calibration and validation sets.
292 For A_n, Ψ, WUE_{intr}, WUE_{inst}, which were phenotyped at low-throughput, the low number of leaves
293 did not allow us to build calibration sets per water treatment. Thus, calibration sets were created
294 including 24 leaves from each water treatment (dataset J in Table 1).
295 The predictive ability of the PLSR models was assessed using validation sets, each including 10
296 leaves per water treatment and accounting for the proportion of plants phenotyped each day in
297 order to be representative of the entire phenotyping series. This sampling was repeated 100
298 times, resulting in 100 distinct calibration and validation sets, in order to provide reliable estimates
299 and low risk of unrepresentative sampling. Significant differences between predictive abilities
300 were assessed with a Tukey HSD post-hoc test, as implemented in R/stats.

Calculation of WUE using predicted A

Intrinsic Water Use Efficiency (WUE_{intr_cal}) and instantaneous Water Use Efficiency (WUE_{inst_cal}) were calculated as the ratios between predicted A_n and the measured g_{sw} or E from the Li600 device, respectively. It was done for each validation set of the dataset J (Table 1). The prediction accuracy was obtained by comparing these calculated WUE_{intr_cal} and WUE_{inst_cal} with WUE_{intr} and WUE_{inst} obtained with the Li6800 measurements. Significance between predictive abilities were assessed with a Tukey HSD post-hoc test, as implemented in R/stats.

Heritability of predicted traits

To evaluate the usefulness of PLSR models we used them to predict phenotypic values across the whole panel. To this aim, a calibration set was built using leaves measured in the "low-throughput" sessions, resulting in 34 leaves per water treatment, and used to train PLSR models (dataset K in Table 1). The evaluation of these models was assessed through their cross-validated statistics R^2_{cv} and $RMSE_{cv}$. Accurate enough models (R^2_{cv} >= 0.55) were then applied to predict traits using high-throughput measurements.

The following mixed-linear model was applied to the resulting predicted values in order to decompose the phenotypic variation using R/lme4 [51]:

$Y_{ij} = \mu + G_i + \Sigma A_{ij} + \varepsilon_{ij}$, where Y_{ij} represented the vector of predicted values per water treatment, μ was the intercept, G_i was the random effect of the cultivar i, ΣA_{ij} was a sum of fixed effects of experimental covariates measured on the plant j from the cultivar i and ε_{ij} was the random residual effect. The fixed effects included date of measurement, time of measurement, time elapsed between the last irrigation of the plant and the measurement, duration between the entrance of the plant into the platform and the measurement, environmental conditions as the average of the 10 probes (temperature, PPFD, relative humidity, vapor pressure deficit), and plant density around each plant. For each trait within each water treatment a model selection was carried to

retain only the significant fixed effects (α = 0.05) using R/rutilstimflutre [52]. As WUE_{intr_cal} was not symmetrically distributed, a square-root transformation was performed before the adjustment of the mixed-linear model. Variance components in the selected model were then estimated and used to estimate broad-sense heritability as: $H^2 = \frac{\sigma^2 G}{\sigma^2 G + (\frac{\sigma^2 E}{n})}$ with $\sigma^2 G$ the genotypic variance, $\sigma^2 E$ the residual variance and n the mean number of repetitions (leaves) per cultivar within each water treatment, typically around 2.

Results

Variability of leaf ecophysiological traits

The lowest coefficient of variation was observed for WC_f (<5%), regardless of the water treatment. LMA and WQ showed intermediate coefficients of variation (between 15% and 22%), with generally slight differences between water treatments and experiments (greenhouse *vs.* outdoors) (Supplementary Table 2). LMA and WQ were positively correlated, with their correlation coefficient increasing from 0.63 to 0.73 with the intensity of water deficit in the greenhouse experiment (Supplementary Figure 1a, Supplementary Figure 1b and Supplementary Figure 1c). LMA and WC_f were negatively correlated, regardless of the water treatment, with correlation coefficient decreasing from -0.57 to -0.48 when water deficit increased. The correlation between WC_f and WQ was lower but remained significant and positively different from zero, with values close to 0.25 (Supplementary Figure 1a, Supplementary Figure 1b and Supplementary Figure 1c).

A_n,Ψ, WUE_{intr} and WUE_{inst} displayed more variability than LMA, WC_f and WQ (Supplementary Figure 2), as underlined by higher coefficients of variation (Supplementary Table 2). A_n was the trait whose coefficient of variation was the most affected by the water treatment, increasing as the water deficit increased. Conversely, the coefficient of variation of WUE_{intr} decreased with

348 increasing water deficit. A_n was not significantly correlated with WUE_{intr} and WUE_{inst}, whatever the
349 water treatment (Supplementary Figure 1d, Supplementary Figure 1e and Supplementary Figure
350 1f). WUE_{intr} was strongly correlated with WUE_{inst} with values around 0.8 in all conditions
351 (Supplementary Figure 1d, Supplementary Figure 1e and Supplementary Figure 1f).

352 Devices complementarity for the prediction of leaf traits

353 Overall, spectra of the three spectrometers collected in the greenhouse experiment (Datasets A
354 and B in Table 1) were highly effective in predicting traits such as LMA, WC_f and WQ. The
355 coefficient of determination computed within the cross-validation procedure consistently
356 exceeded 0.67 (Fig 1a, Supplementary Figure 3a, Supplementary Figure 3b and Supplementary
357 Figure 3c) with almost no variability across the cross-validation repetitions. Nevertheless,
358 significant differences in R^2_{cv} were found between devices (Supplementary Table 3). The ASD
359 spectrometer on dry leaves yielded the most accurate models to predict LMA compared to micro-
360 portable spectrometers used *in situ* (Fig. 1, Supplementary Figure 4a). In contrast, the best
361 models for predicting WC_f and WQ were established with spectra from micro-portable
362 spectrometers collected *in situ* (MicroNIR, NeoSpectra and MicroNIR_NeoSpectra in Fig. 1a).
363 Although model performances were almost similar for these three modalities, slightly better
364 predictions were achieved by combining the wavelengths from MicroNIR and NeoSpectra (Fig.
365 1). The PLSR models built with the variables from the Li600 device were unable to predict LMA,
366 WC_f and WQ ($R^2_{cv} \leq 0.1$, Fig. 1a).
367 Regarding the prediction of A_n, Ψ, WUE_{intr} and WUE_{inst} with spectrometers, the R^2 values remained
368 low and not higher than 0.27 whatever the device considered (Fig. 1b). In contrast, the models
369 built with variables from the Li600 device for predicting A_n displayed an R^2_{cv} of 0.82 and an
370 $NRMSE_{cv}$ of 0.25 (Fig. 1b, Supplementary Figure 4b, Supplementary Figure 3d). The Li600 device
371 also seemed to capture relevant information for predicting WUE_{intr} and WUE_{inst}, with R^2_{cv} values

372 of 0.48 and 0.3, respectively (Fig. 1b, Supplementary Figure 3f and Supplementary Figure 3g).
373 Yet, it is noteworthy that the prediction accuracies of WUE_{inst} using either the Li600 or the
374 MicroNIR_NeoSpectra were not significantly different (Fig. 1b and Supplementary Table 3). Ψ
375 was poorly predicted, regardless of the device used, with an R^2_{cv} always below 0.24 (Fig. 1b and
376 Supplementary Figure 3e). The large variability in R^2_{cv} or $NRMSE_{cv}$ between the different cross-
377 validation sets was attributed to the small number of leaves included in each set (Fig. 1, dataset
378 B in Table 1).
379 To further investigate the effect of calibration set size on model accuracy, we compared PLSR
380 models built with low- (109 leaves) and high-throughput (1437 leaves) measurements of LMA. As
381 expected, models built within the low-throughput calibration set displayed significantly lower R^2_{cv}
382 than models built within the high-throughput calibration set (by 0.08 to 0.15) (Supplementary
383 Figure 5a). This might partly be due to a lower range of variation in the low-throughput calibration
384 set compared to the high-throughput calibration set (Supplementary Figure 1h). This point is
385 confirmed by the $RMSE_{cv}$ values which displayed low differences between low- and high-
386 throughput calibration sets with values always lower than 0.42 mg cm^{-2} (Supplementary Figure
387 5b).

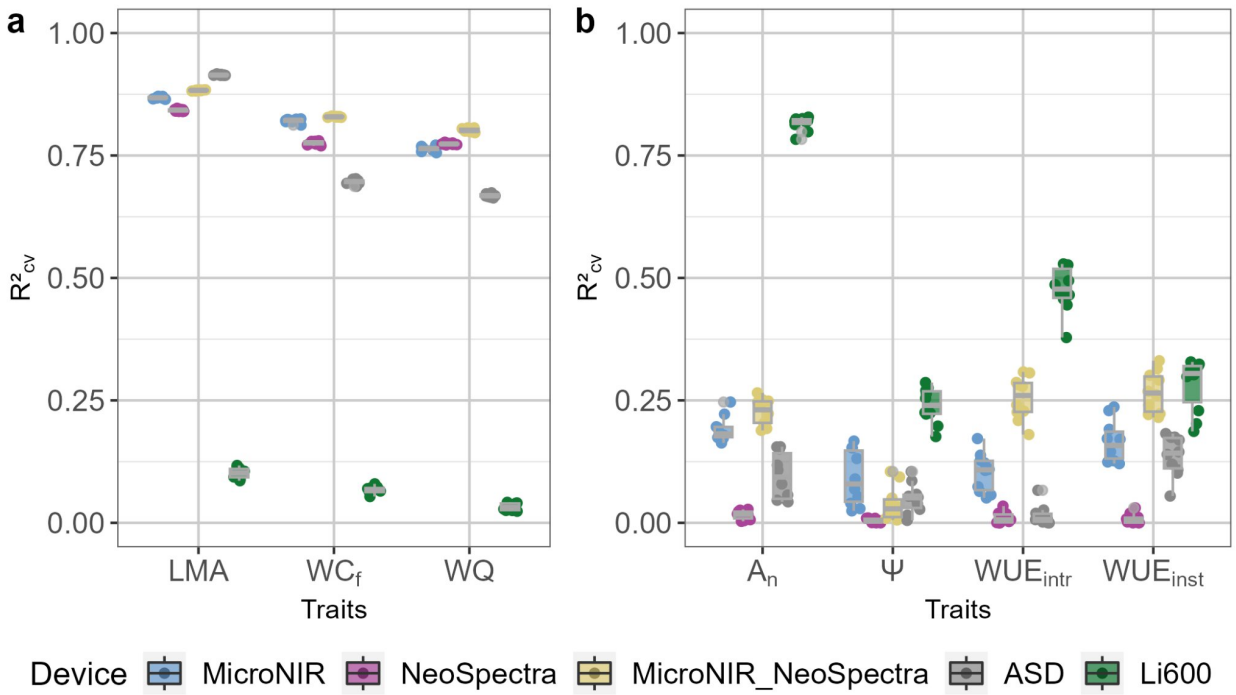

Fig. 1 R^2_{cv} values computed in the cross validation procedure of leaf traits according to the device. **a** High-throughput measured traits : leaf mass per area (LMA), Water content (WC_f) and Water quantity (WQ) (dataset A in Table 1). **b** Low-throughput measured traits : Net CO_2 assimilation (A_n), Leaf water potential (Ψ), Intrinsic water use efficiency (WUE_{intr}) and Instantaneous water use efficiency (WUE_{inst}) (dataset B in Table 1). Each point in the boxplot corresponds to the R^2_{cv} of one of the 10 iterations of the cross-validation (as detailed in Table 1).

Prediction robustness across experiments

The prediction accuracy for LMA (R^2) was significantly lower with the outdoors than the greenhouse validation set (Fig. 2). The median R^2 values of the 15 iterations ranged from 0.30 to 0.86 when models were evaluated with the outdoors validation set and from 0.78 to 0.90 when they were evaluated with the greenhouse validation set when all the devices were considered. The ASD device used on dried disks yielded the most robust models for predicting LMA, with a median R^2 values always exceeding 0.78, and RMSE values lower than 0.62 mg cm^{-2}, regardless of the calibration and validation sets (Fig. 2). Conversely the models built with micro-portable

403 spectrometers on fresh leaves, especially the MicroNIR, appeared to be impacted by the
404 environmental conditions in which they were calibrated and validated (Fig. 2). This was
405 particularly the case when models were calibrated with greenhouse data and validated on data
406 collected outdoors. For MicroNIR, NeoSpectra and MicroNIR_NeoSpectra devices, the
407 performance of models to predict LMA in outdoors conditions was similarly good whether
408 calibrated with outdoor or greenhouse_outdoor calibration sets with R^2 always exceeding 0.61
409 and RMSE always below 0.51 mg cm^{-2} (Fig. 2). Models calibrated with outdoors data remained
410 accurate to predict greenhouse data since R^2 was greater than 0.78 and RMSE lower than 0.8
411 mg cm^{-2}. However, the prediction of greenhouse data, using an outdoors calibration set
412 significantly reduced R^2 and RMSE values compared to greenhouse or greenhouse_outdoor
413 calibration sets (p-value ≤ 0.001). This might partly be due to a lower coefficient of variation for
414 LMA in outdoors (16.7) than in greenhouse (between 20.5 and 21.4) experiments (Supplementary
415 Table 2, Supplementary Figure 2h and Supplementary Figure 2i).

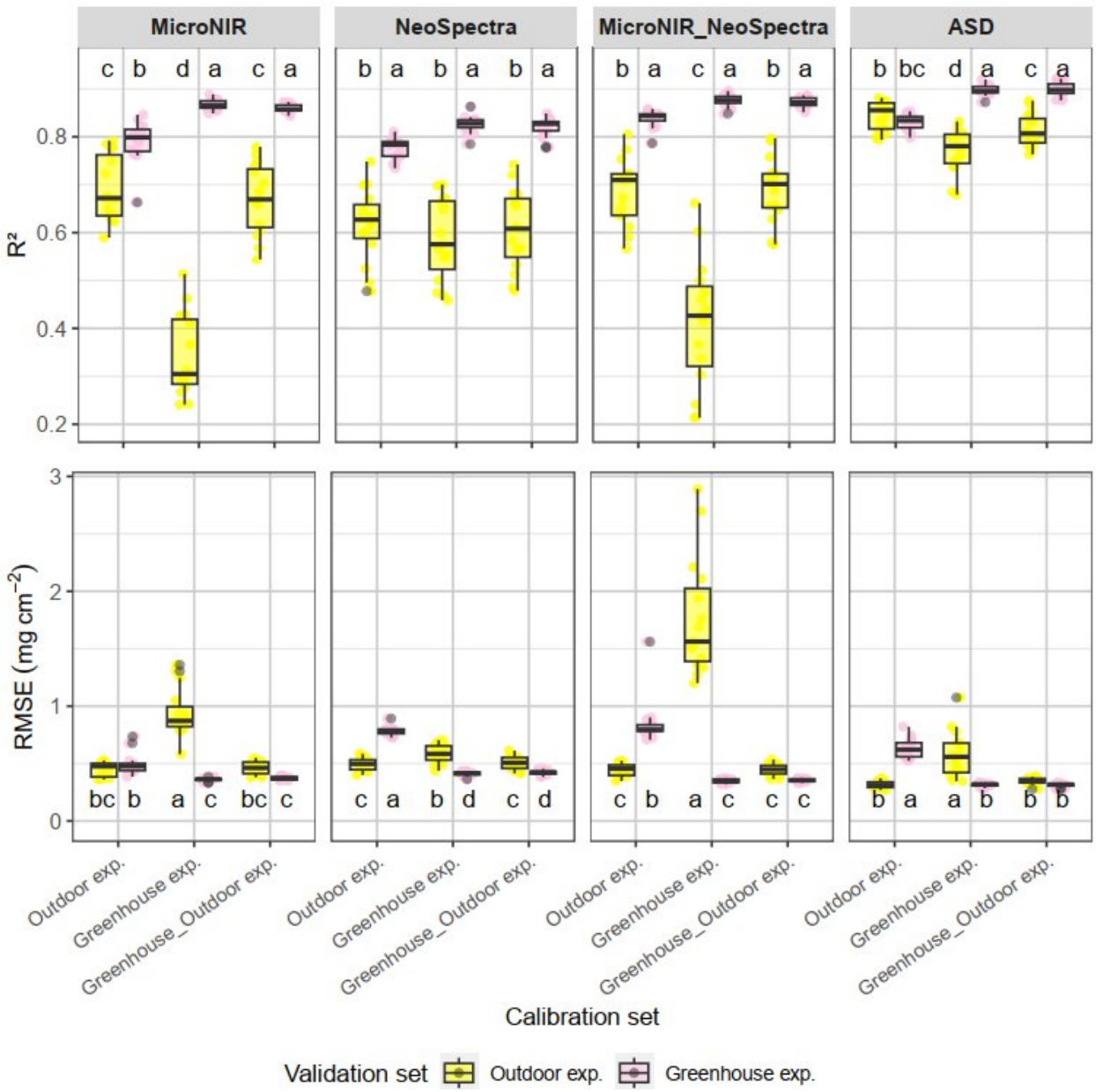

Fig. 2 Prediction quality of LMA depending on whether spectra were collected in outdoors or greenhouse experiments.

The top panel corresponds to the R^2 and the bottom panel corresponds to the RMSE (in mg cm^{-2}) of the prediction. Each point in the boxplot represents one of the 15 iterations of the validation (datasets C, D, E in Table 1). Different calibration sets are highlighted on the x-axis of each panel. Boxplots with the same letters within one facet are not significantly different ($p > 0.05$). exp. = experiment.

423 # Prediction robustness between soil water treatments

424 PLSR models were compared between water treatments for LMA, WC_f and WQ as these 3 traits
425 were available on the entire diversity panel (datasets F to J in Table 1). Partitioning the validation
426 set according to the water treatment had more influence on model performance than partitioning
427 the calibration set (Supplementary Figure 6). We thus decided to focus our analysis on the
428 comparison of PLSR models calibrated with leaves from the 3 water treatments (Fig. 3).
429 In most cases, R^2 of predictive models was impacted by the water treatment (Fig. 3a) while RMSE
430 was not (Fig. 3b). For LMA, R^2 was significantly lower when predicting WD2 compared to WW
431 and WD1 validation sets (Fig. 3a). These differences in R^2 could be related to a decrease in the
432 range of variation with the increase in water deficit (Supplementary Figure 2a), since the
433 corresponding RMSE values did not significantly differ between water treatment (Fig. 3b).
434 Similarly for MicroNIR_NeoSpectra, WC_f was predicted with R^2 that significantly decreased from
435 0.86 to 0.75 with water deficit intensity, while RMSE did not significantly vary (Fig. 3a, Fig. 3b).
436 This influence of water treatment trait prediction was not observed for WQ, with R^2 values always
437 close to 0.8 with MicroNIR_NeoSpectra (Fig. 3a) and median RMSE values lower than 8.1 mg
438 H_2O, regardless of the water treatment considered for the validation sets (Fig. 3b). Spectra from
439 MicroNIR_NeoSpectra, taken *in situ*, provided a robust prediction of WQ across water treatment.
440 On the contrary, the ASD device, used on dried leaves, yielded a less accurate prediction of WQ
441 with median R^2 values ranging from 0.58 to 0.67 depending on the water treatment and RMSE
442 always higher than 9.3 mg H_2O.

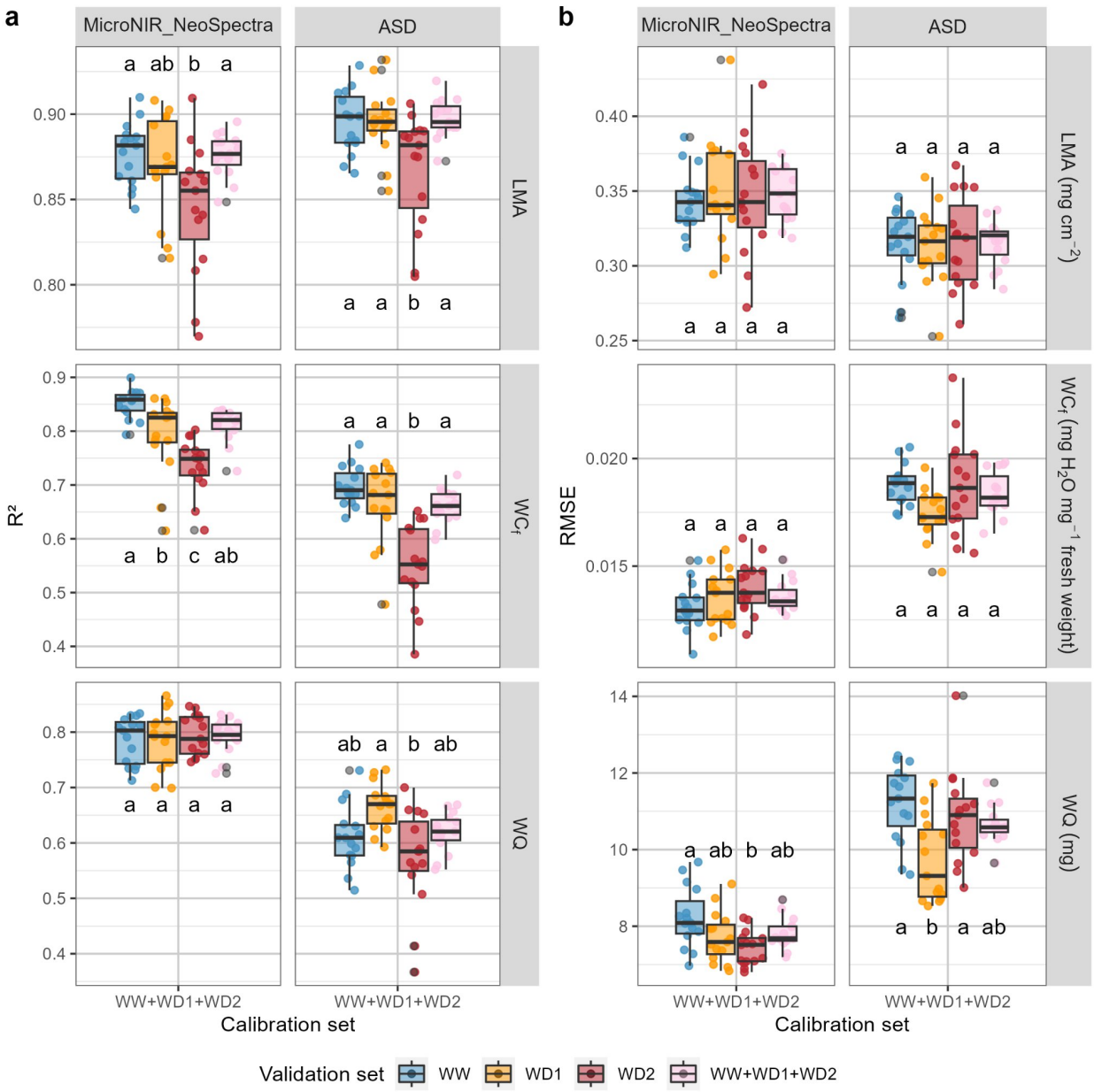

Fig. 3 Prediction quality of PLSR models for LMA, WC_f and WQ depending on the validation set. Models were calibrated using the WW+WD1+WD2 calibration set (dataset I in Table 1) and the MicroNIR_NeoSpectra and ASD devices. **a** corresponds to the R^2 and **b** corresponds to the RMSE in the validation set. Each point in the boxplot represents one of the 15 iterations of the validation (as detailed in Table 1). The blue dots and boxes correspond to the well-watered treatment (WW); orange dots and boxes to the moderate water deficit treatment (WD1); red dots and boxes to the severe water deficit treatment

450 (WD2); and pink dots and boxes to all WW, WD1 and WD2 treatments. LMA = leaf mass per area; WC_f =
451 water content; WQ = water quantity.

452 Because A_n, Ψ, WUE_{intr} and WUE_{inst} were phenotyped on a limited subset of leaves from the
453 diversity panel, PLSR models were only calibrated by gathering all water treatments, and they
454 were evaluated separately for each water treatment (dataset J in Table 1). The high variability of
455 models performance observed with the 100 resampling of validation sets within each water
456 treatment (Fig. 4) was likely a consequence of the small number of leaves in each validation set.
457 A_n was poorly predicted with spectrometers alone, with median R^2 values below 0.37 (Fig. 4a)
458 and median RMSE higher than 3.25 µmol m^{-2} s^{-1} for MicroNIR_NeoSpectra and ASD devices,
459 regardless of the water treatment considered for the validation (Fig. 4, Supplementary Figure 7).
460 In contrast, Li600 measurements resulted in a good prediction of A_n, without much influence of
461 water treatments, with median R^2 higher than 0.72 and median RMSE lower than 1.62 µmol m^{-2}
462 s^{-1} across all validation sets (Fig. 4). Neither the spectrometers nor the Li600 alone could reliably
463 predict Ψ, with median R^2 values consistently lower than 0.21, regardless of the water treatment
464 (Fig. 4a). The median RMSE consistently exceeded 0.33 MPa (Fig. 4b).

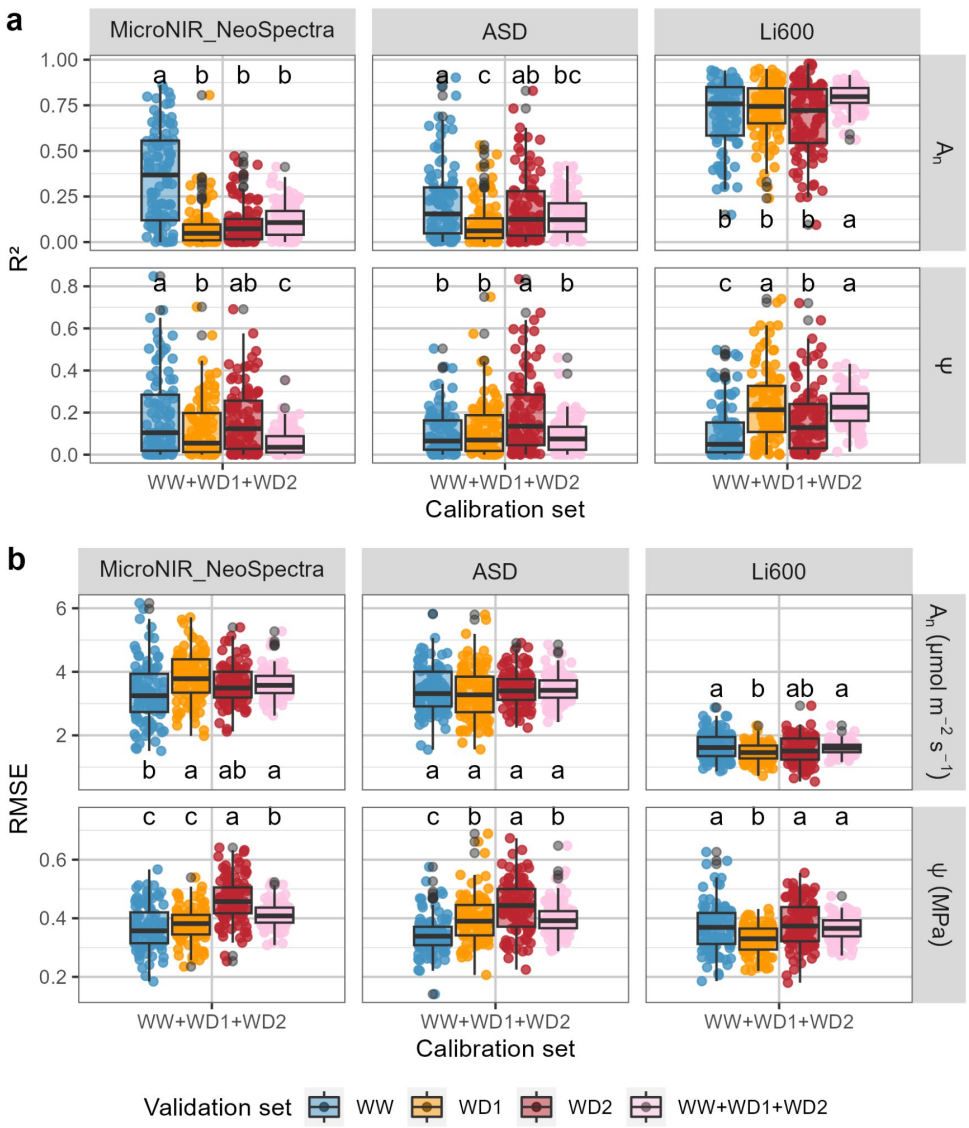

Fig. 4 Prediction quality of PLSR models for A_n and Ψ depending on the validation set. Models were calibrated using the WW+WD1+WD2 calibration set (dataset J in Table 1) and the MicroNIR_NeoSpectra, ASD and Li600 devices. **a** corresponds to the R^2 and **b** corresponds to the RMSE in the validation set. Each point in the boxplot represents one of the 100 iterations of the validation (as detailed in Table 1). The blue dots and boxes correspond to the well-watered treatment (WW); orange dots and boxes to the moderate water deficit treatment (WD1); red dots and boxes to the severe water deficit treatment (WD2); and pink dots and boxes to all WW, WD1 and WD2 treatments. A_n = Net CO_2 assimilation; Ψ = Leaf water potential.

474 PLSR models built with Li600 measurements displayed a variable performance in predicting
475 WUE_{intr} and WUE_{inst} depending on the constitution of the validation set. The validation set that
476 included the three water treatments (WW+WD1+WD2) showed intermediate prediction quality
477 with median R^2 values of 0.52 and 0.46 for predicting WUE_{intr} and WUE_{inst}, respectively (Fig. 5a).
478 The prediction accuracy of PLSR models was significantly different from one water treatment to
479 another. It was maximal under WD1 water treatment for WUE_{intr} and WUE_{inst} with respectively,
480 median R^2 of 0.75 and 0.73, followed by intermediate performance under WW (median R^2 of 0.52
481 and 0.49 for WUE_{intr} and WUE_{inst}) and finally poor performance under WD2 (median R^2 of 0.1 and
482 0.25 for WUE_{intr} and WUE_{inst}) (Fig. 5). These results indicated that PLSR models built with the
483 Li600 device could be considered accurate enough to predict WUE_{intr} and WUE_{inst} for leaves
484 subjected to WD1 water treatment only. On the other hand, spectrometers performed poorly in
485 predicting WUE_{intr} (Supplementary Figure 8) while MicroNIR and MicroNIR_NeoSpectra provided
486 more relevant measurements to predict WUE_{inst} under WD1, with a median R^2 close to 0.53
487 (Supplementary Figure 8a).
488 Since WUE is the ratio between carbon gains (A_n) and water loss related data (g_{sw} or E), both
489 WUE_{intr} and WUE_{inst} could be either directly predicted, or calculated as WUE_{intr_cal} and WUE_{inst_cal}
490 using the prediction of A_n as and measured g_{sw} (or E) from the LI600 device.
491 We first compared values of g_{sw} (or E) directly measured with both devices (Li6800 or Li600). The
492 correlations for g_{sw} and E ranged from 0.93 to 0.97 and from 0.83 to 0.92, respectively, depending
493 on the water treatment (Supplementary Figure 9). Then, we assessed the interest of using
494 WUE_{intr_cal} and WUE_{inst_cal} instead of predicting directly WUE_{intr} and WUE_{inst}. The median R^2 of
495 WUE_{intr_cal} reached 0.77 for the WW validation set which was significantly higher than for the
496 prediction of WUE_{intr} for the same WW validation set (p-value < 0.001) (Fig. 5a). Prediction
497 performance for WUE_{intr_cal} was not significantly improved compared to WUE_{intr} prediction when
498 considering the other validation sets (WD1, WD2) (Fig. 5a and Fig. 5c). The use of calculated

values for both WUE resulted in considerable variation in RMSE values compared with direct predictions for WD1 and WD2 water treatment (Fig. 5b and Fig. 5c).

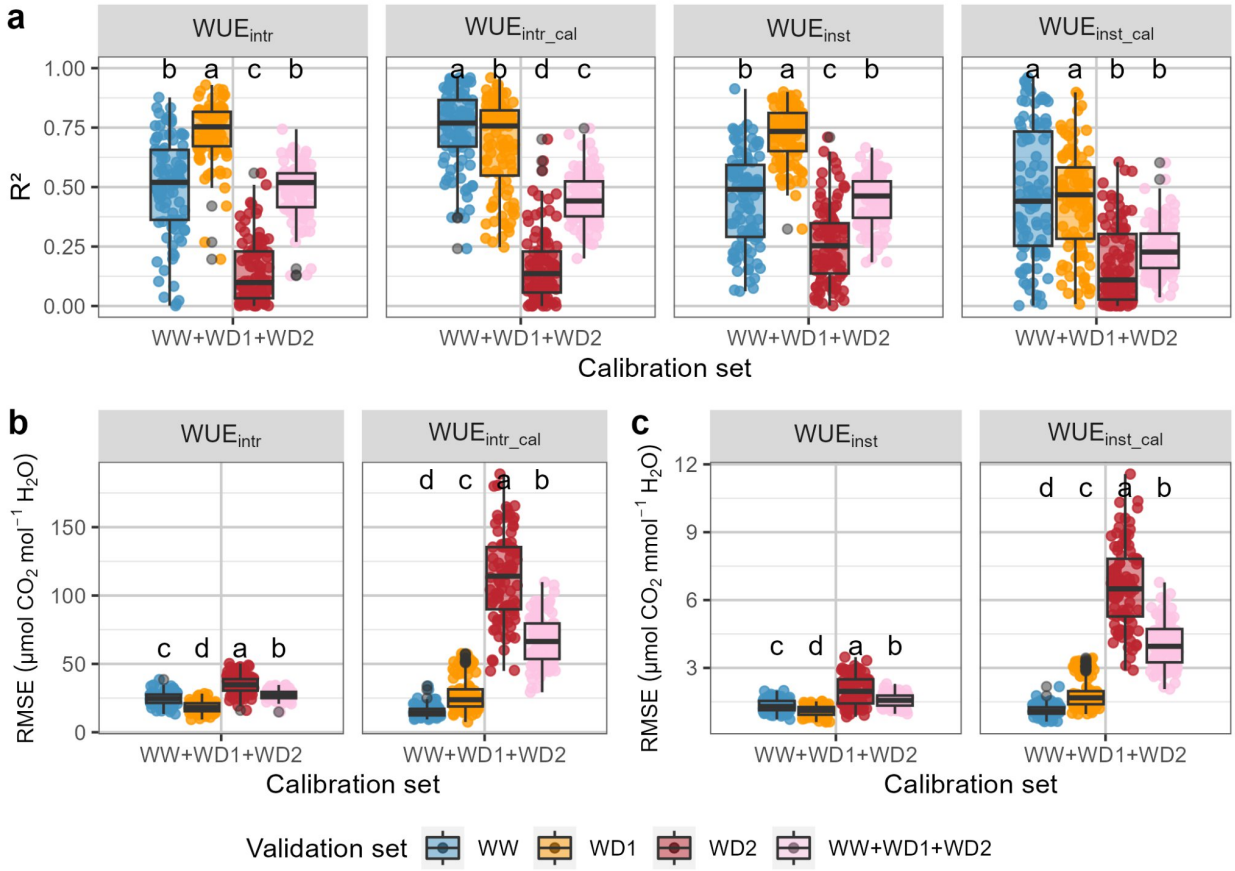

Fig. 5 Prediction quality of PLSR models for WUE_{intr}, WUE_{inst}, WUE_{intr_cal}, WUE_{inst_cal} depending on the validation set.

Models were calibrated with the WW+WD1+WD2 calibration set (dataset J in Table 1) with variables from the Li600 device for intrinsic water use efficiency (WUE_{intr}) and instantaneous water use efficiency (WUE_{inst}) (as the ratio of Li6800 A_n and E or g_{sw} measurements) directly predicted or calculated (WUE_{intr_cal} andWUE_{inst_cal}) as the ratio between (i) the prediction of the net CO_2 assimilation (A_n) and (ii) the Li600 measurements of the stomatal conductance (g_{sw}) and the transpiration rate (E) respectively. **a** corresponds to the R^2, **b** corresponds to the RMSE expressed in µmol CO_2 mol^{-1} H_2O for WUE_{intr} and WUE_{intr_cal} and **c** corresponds to the RMSE expressed in µmol CO_2 $mmol^{-1}$ H_2O for WUE_{inst} and WUE_{inst_cal}. Each point in the boxplot corresponds to one of the 100 iterations of the validation (as detailed in Table 1). The blue dots and

boxes correspond to the well-watered treatment (WW); orange dots and boxes to the moderate water deficit treatment (WD1); and red dots and boxes to the severe water deficit treatment (WD2).

Prediction of LMA, A_n, WUE_{intr} and WUE_{inst} using high-throughput measurements

The PLSR models that were retained to predict traits (LMA, A_n, WUE_{intr} and WUE_{inst}) were those exhibiting a median R^2 higher than 0.55 in at least one validation set when they were built on low throughput data (dataset K in Table 1). The prediction models for LMA used between 3 and 12 latent variables, had R^2_{cv} higher than 0.76 and $RMSE_{cv}$ lower than 0.4 mg cm^{-2}, irrespectively of the water treatment. The cross-validated predicted values of LMA, regardless of the device used, were highly correlated with the *in-situ* measurements, with a correlation coefficient higher than 0.76 (Supplementary Figure 10a). Ten latent variables were retained to predict A_n using the Li600 device (Table 2). The corresponding model displayed good performance for all water treatments with R^2_{cv} and $RMSE_{cv}$ of 0.8 and 1.59 µmol m^{-2} s^{-1} respectively. WUE_{intr} and WUE_{inst} were predicted with an R^2_{cv} of 0.47 and 0.44, across all water treatment. In agreement with previous results, R^2_{cv} for WW and WD2 water treatments did not exceed 0.53 whereas the R^2_{cv} for WD1 was above 0.73 (Table 2, Supplementary Figure 10c and Supplementary Figure 10d). Therefore, models built with Li600 data were used to predict WUE_{intr} and WUE_{inst} for leaves under WD1 treatment only.

The heritability of measured LMA across the entire diversity panel ranged from 0.28 to 0.43 depending on the water treatment (Fig. 6a). The prediction of LMA using spectrometers provided very similar heritability estimates ranging from 0.20 to 0.42. We also estimated genetic values from LMA predictions which showed high correlations ($r \geq 0.9$) with the genetic values estimated with the observed phenotypes, whatever the water treatment considered (Fig. 6c). Predicted A_n, WUE_{intr} and WUE_{inst} had heritability ranging from 0.16 to 0.5, depending on the water treatment

536 (Fig. 6b). WUE_{intr_cal} which was calculated with predicted A_n and measured g_{sw} was unable to
537 provide genetic information with an heritability lower than 0.1. For the WD1 water treatment,
538 WUE_{intr_cal} displayed a lower heritability than WUE_{intr} that was directly predicted with Li600
539 variables (Fig. 6b).

540 **Table 2** PLSR criteria obtained for predictive models (dataset K in Table 1).

			WW+WD1+WD2		WW		WD1		WD2	
Trait	Device	LV	R^2_{cv}	$RMSE_{cv}$	R^2_{cv}	$RMSE_{cv}$	R^2_{cv}	$RMSE_{cv}$	R^2_{cv}	$RMSE_{cv}$
LMA	MicroNIR	12	0.78	0.38	0.83	0.40	0.82	0.33	0.58	0.42
LMA	NeoSpectra	3	0.76	0.40	0.84	0.41	0.74	0.40	0.66	0.38
LMA	MicroNIR_ NeoSpectra	4	0.82	0.34	0.88	0.35	0.86	0.30	0.64	0.39
LMA	ASD	8	0.83	0.34	0.84	0.38	0.80	0.36	0.80	0.27
A_n	Li600	10	0.80	1.59	0.74	1.63	0.74	1.57	0.69	1.56
WUE_{intr}	Li600	8	0.48	27.09	0.53	22.71	0.73	17.90	0.02	36.96
WUE_{inst}	Li600	7	0.44	1.53	0.46	1.25	0.74	1.01	0.17	2.11

541

542 PLSR models follows a cross-validation procedure using leaves directly measured with conventional
543 methods to predict Leaf mass per area (LMA), Net CO_2 assimilation (A_n), Intrinsic water use efficiency
544 (WUE_{intr}) and Instantaneous water use efficiency (WUE_{inst}) on the diversity panel.
545 LV corresponds to the optimal number of latent variables retained by the model. R^2_{cv} and $RMSE_{cv}$ are
546 calculated for all water treatment together (102 leaves) and for each water treatment separately (34 leaves):
547 Well-watered treatment (WW), Moderate water deficit treatment (WD1), Severe water deficit treatment
548 (WD2).

549

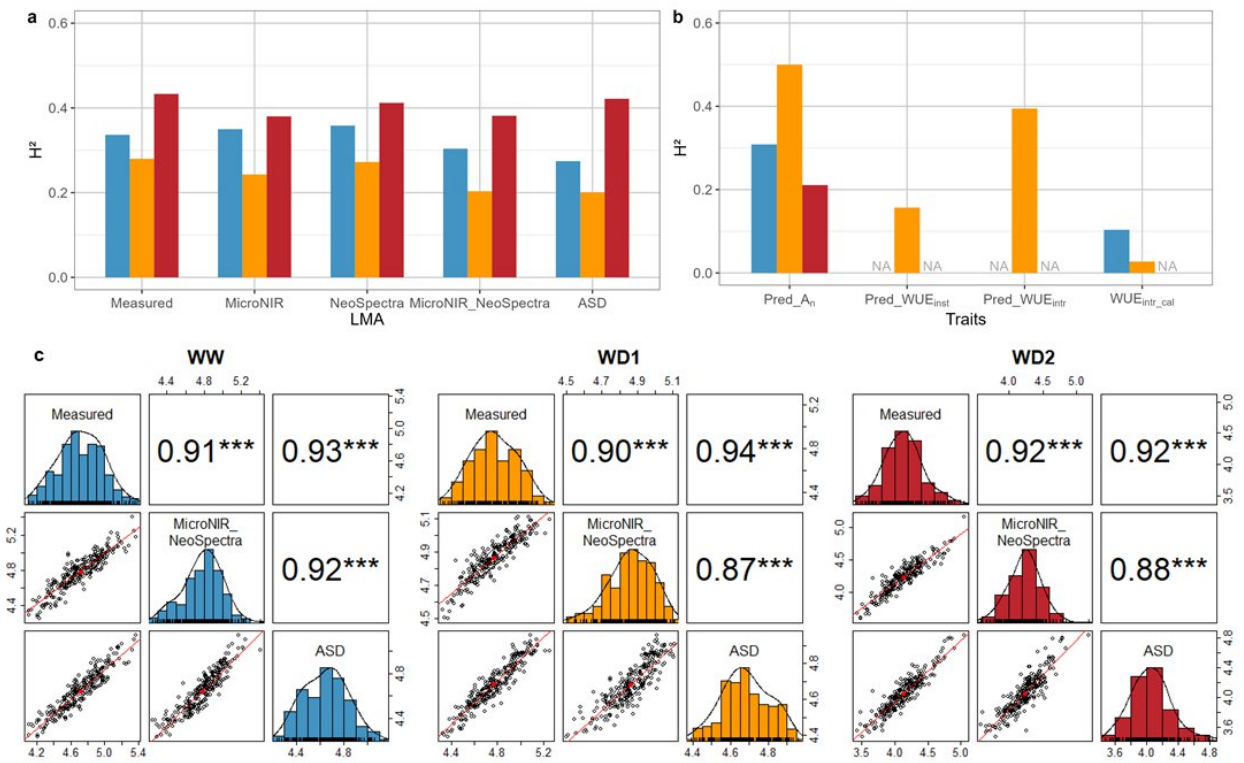

Fig. 6 Estimation of the heritability of observed and predicted leaf traits under different water treatments.

a Corresponds to the broad-sense heritability estimates for measured and predicted leaf mass per area (LMA) according to the spectrometer used as a predictor; **b** corresponds to the heritability estimates for predicted net CO_2 assimilation (Pred_A_n), predicted instantaneous water use efficiency (Pred_WUE_{inst}), predicted intrinsic water use efficiency (Pred_WUE_{intr}) and the square-root of the calculated intrinsic water use efficiency (WUE_{intr_cal}); **c** corresponds to the distributions and correlations between genotypic values of measured LMA *versus* genotypic values of predicted LMA according to the spectrometer used as a predictor (MicroNIR_NeoSpectra or ASD). Blue bars represent the well-watered treatment (WW); yellow bars represent the moderate water-deficit treatment (WD1); red bars represent the severe water deficit treatment (WD2). NA indicates that the trait was not predicted (or calculated) under the corresponding water treatment. *** = p-value ≤ 0.001.

Discussion

In this study, we explored the performance of models built with high-throughput devices to predict ecophysiological and morphological leaf traits across a large genetic diversity panel of *Vitis vinifera* varieties. Using a combination of large datasets collected on thousands of grapevine plants across different experiments and conditions, we were able to evaluate the robustness of predictive models with respect to the experimental conditions. Based on an analytical framework with dedicated and repeated calibration and validation data sets, we found that the prediction accuracy within the validation set containing leaves of plants grown under severe water deficit was generally lower than that for other water treatments. We also found that all devices (NIR spectrometers or poro-fluorometer) were efficient in predicting at least one trait. Finally, we found that the leaf traits predicted with NIRS and poro-fluorometry data captured a substantial part of the genetic variation which paves the way towards their use in genetic studies.

High-throughput devices are complementary to predict a range of morphological and physiological traits

For too small datasets or too complex traits, repeated double cross-validation is more efficient to obtain an accurate estimation of the prediction error compared to a simple cross-validation [39]. In our analyses, we performed 100 resampling iterations for each cross-validated calibration and validation sets for predicting A_n, Ψ, WUE_{intr} and WUE_{inst}. Previous studies most often only made one iteration for cross-validating PLSR models [25]. In this study, we observed that model R^2 was highly dependent on the iteration. Consequently, depending on the sampling for unique datasets of calibration and validation, the models could have been biased. Spectrometers resulted in better performances than Li600 for the prediction of LMA, WC_f and WQ while the Li600 was more efficient for the prediction of A_n, WUE_{intr} and WUE_{inst}. The laboratory spectrometer (ASD) was found as the most accurate to predict LMA. Absorption bands between 1670-1830 nm and 2000-2200 nm, covered by this device, were previously reported as relevant

588 for predicting leaf dry matter related traits like protein, lignin and cellulose contents [53–56]. The
589 difference of predictive performance between the laboratory spectrometer and the portable
590 spectrometers could also be due to the presence/absence of water in the corresponding leaf
591 samples as spectra from the ASD device were collected on dried leaves while spectra from
592 MicroNIR and NeoSpectra devices were collected *in situ*. It is known that the presence of water
593 in samples can dominate the signal, potentially reducing the detection of smaller peaks related to
594 other components such as nitrogen [33], protein, lignin and cellulose, which could be essential
595 for predicting LMA. LMA and WC_f were negatively correlated, likely because leaves with a low
596 LMA tend to have a high water content [57], but we observed that the devices yielding the best
597 PLSR models for predicting these traits differed. Water-related traits (WC_f and WQ) were better
598 predicted *in situ* using portable spectrometers while LMA was better predicted *ex situ* using a lab
599 spectrometer on dried leaves. Moreover, water-related traits were better predicted using the
600 combined wavelengths of MicroNIR and NeoSpectra than with either device alone. This is likely
601 because their combined spectral range includes the two main water absorption bands around
602 1400–1440 nm and 1900–1950 nm [29].

603 The poro-fluorometer device was found to provide good prediction of A_n. To our
604 knowledge, this is the first report of A_n prediction using rapid and simultaneous porometry and
605 fluorescence measurements. This aligns with previous studies showing that A_n correlates with
606 the fluorescence decline ratio under steady-state conditions with saturated light [58]. Additionally,
607 A_n was well predicted when combining the electron transport rate, also derived from leaf
608 fluorescence measurements, with the leaf-to-air temperature difference, used as a proxy for
609 stomatal closure [14,16,17]. In our study, spectrometers could not accurately predict A_n ,
610 conversely to other studies on maize [28]. At least 3 spectra from the same side of a leaf were
611 used in the maize experiment, while in our study, two spectra (one on each side) were taken from
612 the same spatial area of a given leaf. De Bei et al. [29] reported that decreasing the number of
613 spectra collected on a given leaf from 6 to 1 decreases the prediction accuracy. However,

614 increasing the number of spectra increases the phenotyping time which is not suitable for high-
615 throughput phenotyping.

616 We found that the Li600 device was able to predict WUE_{intr} and WUE_{inst} for plants under
617 a moderate water deficit. In contrast, spectral data generally failed to provide reliable predictions,
618 except for WUE_{inst}, which could be predicted using MicroNIR and the MicroNIR-NeoSpectra
619 combination under moderate water deficit. With a larger sample size, Cotrozzi et al. [28] reported
620 fairly good predictive performance of models built with NIRS data by performing 500 randomized
621 permutations for WUE_{intr} and WUE_{inst} on an external validation set of maize.

622 The leaf water potential is a physiological trait commonly used as an indicator of soil water
623 deficit intensity [3,29,59]. In our study, the leaf water potential was poorly predicted, regardless
624 of the device used. Experiments on maize have shown that the leaf water potential could be
625 predicted by NIRS with a R^2_{cv} of 0.63 on both greenhouse and field samples [28]. Field
626 experiments on grapevines have shown the interest of NIRS for the prediction of leaf water
627 potential on Syrah and Cabernet Sauvignon [29]. However, in these studies, stem water potential
628 was better predicted than leaf water potential perhaps because spectra, taken at a single spot on
629 the leaf, are not representative of the spatial variability in stomatal conductance [29,60]. Moreover,
630 in the case of a slow soil drying, the osmotic adjustment leads to a maintenance in water content
631 even when leaf water potential decreases [3] which could explain the good predictive ability of
632 water content compared to leaf water potential.

633 Recently, the PLSR method was extended to the combination of multiple methods and
634 sources of high-throughput data using deep learning [61] or multi-block techniques such as SO-
635 PLS or ROSA [62,63]. The prediction of amylose, starch and cellulose contents, using a SO-PLS
636 model with NIR and mid-infrared spectra, had a better R^2 and a smaller error compared to the
637 PLSR models based on a NIR and mid-infrared spectra separately [64]. The prediction of E using
638 NIRS data was shown to be improved by incorporating climatic data such as radiation,
639 temperature, relative humidity and wind speed [38]. It has also recently been shown that the

640 prediction of leaf water potential could be improved by combining leaf spectral reflectance and
641 leaf orientation parameters [65]. However, in a high-throughput phenotyping approach, some
642 complementary information such as leaf orientation, might be difficult to phenotype at a large
643 scale.

644 The robustness of PLSR models depends on the traits and the 645 plant growing conditions

646 LMA has been widely used across species and locations to investigate the robustness of
647 predictive models when plants were subjected to different environmental conditions [26,66]. In
648 these studies, a considerable decrease in model accuracy was found when a trained PLSR model
649 was applied to a validation set containing other sites or experiments. Here, we consistently
650 highlighted a degradation of outdoors prediction using a greenhouse calibration set, similar to
651 recent reports on maize [28]. By contrast, a previous interspecific study using more than 700
652 species, reported that the prediction accuracy for LMA could reach a satisfying stable level with
653 a certain number of sites used in the calibration set [66].

654 Furthermore, here we aimed at deciphering the effect of the watering conditions on the
655 PLSR models prediction accuracy. The prediction of ecophysiological traits might indeed be
656 sensible to soil water availability. While most published studies pooled together data collected in
657 different growing conditions in their calibration and validation datasets [29,38], here we tested
658 both the effects of calibrating PLSR models per water treatment, and more importantly of
659 validating them under different water treatments. Our results highlighted that the composition of
660 the validation dataset is more influential than the calibration one. Validating PLSR models per
661 water treatment prevents an artificial inflation of R^2 value caused by the inclusion of multiple water
662 treatments which in turn increases the trait variance. This likely led to more accurate PLSR model
663 evaluations which highlighted changes in prediction accuracy for LMA, WC_f, WUE_{intr} and WUE_{inst}

664 depending on the water treatments. Such variation in R^2 is primarily due to the reduction in trait
665 variance as underlined by the smaller differences observed for RMSE across water treatments.
666 Yet, differences in RMSE remained and they could be attributed to a differential implication of
667 elementary processes across water treatments, some of them being more or less captured by
668 high-throughput measurements.

669 PLSR models allow capturing the genetic variability of leaf traits

670 High correlations between predicted and measured values do not necessarily imply an accurate
671 prediction of genetic variability for the traits, especially when the calibration set includes multiple
672 experimental conditions. In such cases, the model might have been trained to predict differences
673 between experimental treatments rather than between plant-related variation within a given
674 treatment.

675 Our results demonstrated that the broad sense heritability of predicted LMA was
676 comparable to that of measured LMA, underlying, at least for this trait, that predicted values were
677 as efficient as measured values for capturing the genetic variability within each water treatment.
678 Similar findings were reported for chlorophyll content, nitrogen content and specific leaf area,
679 predicted using hyperspectral reflectance data in maize grown under low and high nitrogen
680 conditions [67]. The heritability of predicted A_n, WUE_{intr} and WUE_{inst} was water treatment
681 dependent with a higher heritability for A_n and WUE_{intr} under moderate water deficit compared to
682 WUE_{inst} and other water treatments. One reason to explain these results could be that our study
683 was based on a single leaf per plant even though the choice of the leaf was homogeneous
684 between plants and chosen as well-exposed and fully expanded. Even so, we could not take into
685 account the intra-plant variability between well-exposed or shadowed leaves, their age or even
686 the plant architecture. Increasing the number of measurements is likely to increase trait heritability
687 [68], but this will also extend the phenotyping time, which may not be appropriate in a high-

688 throughput phenotyping strategy. To overcome this problem, other methodologies based on
689 contactless devices could be tested, such as hyper-spectral imaging [69]. Thermal imaging could
690 also be used to assess leaves and canopy temperature as proxies of water stress responses [70].
691 Architectural and functional traits have successfully been characterized using LiDAR,
692 multispectral and thermal imaging in a large core-collection of apple trees [17], highlighting a
693 strong genotypic effect on all traits, with high broad-sense heritability. A significant effect of the
694 water treatment was also detected on most of the traits. However, when traits exhibit high spatial
695 variability at fine scale, the accuracy of phenotyping using imaging sensors is related to the image
696 resolution. Airborne imaging may also be affected by background noise and atmospheric
697 scattering, which can alter the light absorption features of pigments [71]. Nevertheless, these
698 high-throughput methods are highly effective for phenotyping large numbers of plants and multiple
699 environments in a short time, thus reducing the error associated with the plant circadian cycle.
700 These phenotyping methods open prospects for further analyses of the genetic determinism of
701 ecophysiological traits as recently shown in apple trees [72].
702 Overall our results open promising avenues for further deploying genome-wide
703 association studies on traits predicted from high-throughput measurements.

704 Conclusion

705 Our findings demonstrate the strengths of high-throughput devices for predicting ecophysiological
706 traits not directly accessible at the same throughput in a diversity panel of grapevine under varying
707 experimental conditions. We have highlighted the importance of validating PLSR models for each
708 experimental condition and not being restricted to a single calibration/validation set. This
709 approach can be applied to other abiotic stresses, other populations or even other species. This
710 would require sharing high-throughput dataset collected on multiple species, growth conditions
711 and environments to increase predictive models robustness so that they can be applied more

easily to a wide range of study cases. High-throughput phenotyping strategies can be integrated with imaging-based phenotyping or 3D modeling to improve the assessment of inter-plant variability and further enhance our understanding of plant responses to environmental stresses through genetic analyses.

Availability of data and materials

The datasets supporting the conclusions of this article are available in the Recherche Data Gouv repository at https://doi.org/10.57745/WVAPOL. The scripts for prediction analyses are publicly available in GitLab at https://forgemia.inra.fr/eva.coindre/robustness_ht_prediction_ecophysiological_traits.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the ANR (French Research Agency) through the G²WAS project (ANR-19CE20-0024) and the #Digitag project (ANR-16-CONV-0004). PhenoArch platform is funded by the project ANR-464 11-INBS-0012 (PHENOME-EMPHASIS). This work was publicly

funded through the Occitanie Region's program "Défi clé Vinid'Occ" run by the University of Montpellier (PhD grant to EC within the project PlastiVigne).

Authors' contributions

AC-L, VS designed the project and managed the outdoors experiment. TS, AD, AC-L, VS managed the greenhouse experiment in the frame of the G²WAS project. ML and LC-B managed the PhenoArch platform. RB, LC, VF, TL, VB, EC, TS, AC-L and VS collected data. RB, LC and EC curated data. LC and VF performed preliminary analyses. EC analyzed data with input from MR, BP, AC-L and VS. AC-L, VS and BP supervised research. EC, AC-L, BP and VS wrote the manuscript with input from all authors.

Acknowledgments

We are grateful to the INRAE "Domaine de Vassal" Grape Germplasm Repository for collection maintenance. We thank Freddy Gavanon and Mélyne Falcon for the plant management and the preparation of the greenhouse experiment. We thank Gaëlle Rolland for her help in the leaf phenotyping. We thank Martin Ecarnot and Frederic Compan for the access to the laboratory spectrometer ASD. We thank the ETAP (LEPSE), M3P (LEPSE) and DAAV (AGAP Institut) team members for providing help during the preparation and phenotyping of the grapevine panel. We thank Laurent Torregrosa for administration of ANR G²WAS project and Dominique This and Roberto Bacilieri for administration of Défi clé Vinid'Occ PlastiVigne.

References

1. Semia Cherif, Doblas-Miranda E, Lionello P, Borrego C, Giorgi F, Iglesias A, et al. First Mediterranean Assessment Report - Chapter 2: Drivers of Change. Zenodo; 2020 Nov.

752 2. Van Leeuwen C, Sgubin G, Bois B, Ollat N, Swingedouw D, Zito S, et al. Climate change
753 impacts and adaptations of wine production. Nat Rev Earth Environ. 2024;5:258–75.

754 3. Simonneau T, Lebon E, Coupel-Ledru A, Marguerit E, Rossdeutsch L, Ollat N. Adapting plant
755 material to face water stress in vineyards: which physiological targets for an optimal control of
756 plant water status? OENO One. 2017;51:167.

757 4. Damour G, Simonneau T, Cochard H, Urban L. An overview of models of stomatal
758 conductance at the leaf level: Models of stomatal conductance. Plant, Cell & Environment.
759 2010;no-no.

760 5. Bhattacharya A. Effect of Soil Water Deficit on Growth and Development of Plants: A Review.
761 Soil Water Deficit and Physiological Issues in Plants. Singapore: Springer Singapore; 2021. p.
762 393–488.

763 6. Coupel-Ledru A. Plant Water-Use Efficiency. In: John Wiley & Sons, Ltd, editor. eLS. 1st ed.
764 Wiley; 2021. p. 1–8.

765 7. Bota BJ, Flexas J, Medrano H. Genetic variability of photosynthesis and water use in Balearic
766 grapevine cultivars. Annals of Applied Biology. 2001;138:353–61.

767 8. Tomás M, Medrano H, Escalona JM, Martorell S, Pou A, Ribas-Carbó M, et al. Variability of
768 water use efficiency in grapevines. Environmental and Experimental Botany. 2014;103:148–57.

769 9. Prieto JA, Lebon É, Ojeda H. Stomatal behavior of different grapevine cultivars in response to
770 soil water status and air water vapor pressure deficit. OENO One. 2010;44:9.

771 10. Coupel-Ledru A, Lebon É, Christophe A, Doligez A, Cabrera-Bosquet L, Péchier P, et al.
772 Genetic variation in a grapevine progeny (Vitis vinifera L. cvs Grenache×Syrah) reveals
773 inconsistencies between maintenance of daytime leaf water potential and response of
774 transpiration rate under drought. Journal of Experimental Botany. 2014;65:6205–18.

775 11. Trenti M, Lorenzi S, Bianchedi PL, Grossi D, Failla O, Grando MS, et al. Candidate genes
776 and SNPs associated with stomatal conductance under drought stress in Vitis. BMC Plant Biol.
777 2021;21:7.

778 12. Xiao Q, Bai X, Zhang C, He Y. Advanced high-throughput plant phenotyping techniques for
779 genome-wide association studies: A review. Journal of Advanced Research. 2022;35:215–30.

780 13. Murchie EH, Lawson T. Chlorophyll fluorescence analysis: a guide to good practice and
781 understanding some new applications. Journal of Experimental Botany. 2013;64:3983–98.

782 14. Han J, Chang CY, Gu L, Zhang Y, Meeker EW, Magney TS, et al. The physiological basis
783 for estimating photosynthesis from Chl *a* fluorescence. New Phytologist. 2022;234:1206–19.

784 15. Fernández-Calleja M, Monteagudo A, Casas AM, Boutin C, Pin PA, Morales F, et al. Rapid
785 On-Site Phenotyping via Field Fluorimeter Detects Differences in Photosynthetic Performance in
786 a Hybrid—Parent Barley Germplasm Set. Sensors. 2020;20:1486.

16. Losciale P, Manfrini L, Morandi B, Pierpaoli E, Zibordi M, Stellacci AM, et al. A multivariate approach for assessing leaf photo-assimilation performance using the I_{PL} index. Physiologia Plantarum. 2015;154:609–20.

17. Coupel-Ledru A, Pallas B, Delalande M, Boudon F, Carrié E, Martinez S, et al. Multi-scale high-throughput phenotyping of apple architectural and functional traits in orchard reveals genotypic variability under contrasted watering regimes. Hortic Res. 2019;6:52.

18. Foley WJ, McIlwee A, Lawler I, Aragones L, Woolnough AP, Berding N. Ecological applications of near infrared reflectance spectroscopy - a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. Oecologia. 1998;116:293–305.

19. Petisco C, García-Criado B, Mediavilla S, Vázquez De Aldana BR, Zabalgogeazcoa I, García-Ciudad A. Near-infrared reflectance spectroscopy as a fast and non-destructive tool to predict foliar organic constituents of several woody species. Anal Bioanal Chem. 2006;386:1823–33.

20. Cabrera-Bosquet L, Crossa J, Von Zitzewitz J, Serret MD, Luis Araus J. High-throughput Phenotyping and Genomic Selection: The Frontiers of Crop Breeding Converge^{F}. JIPB. 2012;54:312–20.

21. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2001;58:109–30.

22. Ge Y, Atefi A, Zhang H, Miao C, Ramamurthy RK, Sigmon B, et al. High-throughput analysis of leaf physiological and chemical traits with VIS–NIR–SWIR spectroscopy: a case study with a maize diversity panel. Plant Methods. 2019;15:66.

23. De Bei R, Fuentes S, Sullivan W, Edwards EJ, Tyerman S, Cozzolino D. Rapid measurement of total non-structural carbohydrate concentration in grapevine trunk and leaf tissues using near infrared spectroscopy. Computers and Electronics in Agriculture. 2017;136:176–83.

24. Ely KS, Burnett AC, Lieberman-Cribbin W, Serbin SP, Rogers A. Spectroscopy can predict key leaf traits associated with source–sink balance and carbon–nitrogen status. Journal of Experimental Botany. 2019;70:1789–99.

25. Van Wyngaard E, Blancquaert E, Nieuwoudt H, Aleixandre-Tudo JL. Infrared Spectroscopy and Chemometric Applications for the Qualitative and Quantitative Investigation of Grapevine Organs. Front Plant Sci. 2021;12:723247.

26. Serbin SP, Wu J, Ely KS, Kruger EL, Townsend PA, Meng R, et al. From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. New Phytologist. 2019;224:1557–68.

27. Silva-Perez V, Molero G, Serbin SP, Condon AG, Reynolds MP, Furbank RT, et al. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. Journal of Experimental Botany. 2018;69:483–96.

28. Cotrozzi L, Peron R, Tuinstra MR, Mickelbart MV, Couture JJ. Spectral Phenotyping of Physiological and Anatomical Leaf Traits Related with Maize Water Status. Plant Physiol. 2020;184:1363–77.

29. De Bei R, Cozzolino D, Sullivan W, Cynkar W, Fuentes S, Dambergs R, et al. Non-destructive measurement of grapevine water potential using near infrared spectroscopy: Measure of grapevine water potential using NIR. Australian Journal of Grape and Wine Research. 2011;17:62–71.

30. Rapaport T, Hochberg U, Shoshany M, Karnieli A, Rachmilevitch S. Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. ISPRS Journal of Photogrammetry and Remote Sensing. 2015;109:88–97.

31. Workman Jr. J, Weyer L. Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy. 0 ed. CRC Press; 2012.

32. Das B, Sahoo RN, Pargal S, Krishna G, Verma R, Viswanathan C, et al. Evaluation of different water absorption bands, indices and multivariate models for water-deficit stress monitoring in rice using visible-near infrared spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2021;247:119104.

33. Masemola C, Cho MA. Estimating leaf nitrogen concentration from similarities in fresh and dry leaf spectral bands using a model population analysis framework. International Journal of Remote Sensing. 2019;40:6841–60.

34. Ramoelo A, Skidmore AK, Schlerf M, Mathieu R, Heitkönig IMA. Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. ISPRS Journal of Photogrammetry and Remote Sensing. 2011;66:408–17.

35. Roger J-M, Chauchard F, Bellon-Maurel V. EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemometrics and Intelligent Laboratory Systems. 2003;66:191–204.

36. Zhu C, Fu X, Zhang J, Qin K, Wu C. Review of portable near infrared spectrometers: Current status and new techniques. Journal of Near Infrared Spectroscopy. 2022;30:51–66.

37. Barthès BG, Kouakoua E, Clairotte M, Lallemand J, Chapuis-Lardy L, Rabenarivo M, et al. Performance comparison between a miniaturized and a conventional near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen. Geoderma. 2019;338:422–9.

38. Ryckewaert M, Héran D, Simonneau T, Abdelghafour F, Boulord R, Saurin N, et al. Physiological variable predictions using VIS–NIR spectroscopy for water stress detection on grapevine: Interest in combining climate data using multiblock method. Computers and Electronics in Agriculture. 2022;197:106973.

39. Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. Journal of Chemometrics. 2009;23:160–71.

40. Nicolas SD, Péros J-P, Lacombe T, Launay A, Le Paslier M-C, Bérard A, et al. Genetic diversity, linkage disequilibrium and power of a large grapevine (Vitis vinifera L) diversity panel newly designed for association studies. BMC Plant Biol. 2016;16:74.

41. Coupel-Ledru A, Westgeest AJ, Albasha R, Millan M, Pallas B, Doligez A, et al. Clusters of grapevine genes for a burning world. New Phytologist. 2024;nph.19540.

42. Cabrera-Bosquet L, Fournier C, Brichet N, Welcker C, Suard B, Tardieu F. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. New Phytologist. 2016;212:269–81.

43. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2023. Available from: https://www.R-project.org/

44. Brandolini-Bunlon M, Jaillais B, Roger J-M, Lesnoff M. rchemo: Dimension Reduction, Regression and Discrimination for Chemometrics [Internet]. 2023 [cited 2024 Oct 9]. p. 0.1-3. Available from: https://CRAN.R-project.org/package=rchemo

45. Ecarnot M, Compan F, Roumet P. Assessing leaf nitrogen content and leaf mass per unit area of wheat in the field throughout plant cycle with a portable spectrometer. Field Crops Research. 2013;140:44–50.

46. Barnes RJ, Dhanoa MS, Lister SJ. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. Appl Spectrosc. 1989;43:772–7.

47. signal developers. signal: Signal processing [Internet]. 2023. Available from: https://r-forge.r-project.org/projects/signal/

48. Savitzky Abraham, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal Chem. 1964;36:1627–39.

49. Liland KH, Mevik B-H, Wehrens R. pls: Partial Least Squares and Principal Component Regression. 1999. p. 2.8-5.

50. Wong T-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognition. 2015;48:2839–46.

51. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using **lme4**. J Stat Soft. 2015;67.

52. Flutre T, Chabrault. timflutre/rutilstimflutre: v0.170.0 [Internet]. Zenodo; 2019. Available from: https://zenodo.org/record/3580267

53. Curran PJ. Remote sensing of foliar chemistry. Remote Sensing of Environment. 1989;30:271–8.

54. Kokaly RF. Investigating a Physical Basis for Spectroscopic Estimates of Leaf Nitrogen Concentration. Remote Sensing of Environment. 2001;75:153–61.

55. Kawamura K, Watanabe N, Sakanoue S, Inoue Y. Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection. Grassland Science. 2008;54:131–45.

56. Cheng T, Rivard B, Sánchez-Azofeifa A. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sensing of Environment. 2011;115:659–70.

57. Poorter H, Niinemets Ü, Poorter L, Wright IJ, Villar R. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. New Phytologist. 2009;182:565–88.

58. Lichtenthaler HK, Buschmann C, Knapp M. How to correctly determine the different chlorophyll fluorescence parameters and the chlorophyll fluorescence decrease ratio R_{Fd} of leaves with the PAM fluorometer. Photosynt. 2005;43:379–93.

59. Girona J, Mata M, Del Campo J, Arbonés A, Bartra E, Marsal J. The use of midday leaf water potential for scheduling deficit irrigation in vineyards. Irrig Sci. 2006;24:115–27.

60. Albasha R, Fournier C, Pradal C, Chelle M, Prieto JA, Louarn G, et al. HydroShoot: a functional-structural plant model for simulating hydraulic structure, gas and energy exchange dynamics of complex plant canopies under water deficit—application to grapevine (*Vitis vinifera*). in silico Plants. 2019;1:diz007.

61. Vaillant A, Beurier G, Cornet D, Rouan L, Vile D, Violle C, et al. NIRSpredict: a platform for predicting plant traits from near infra-red spectroscopy. BMC Plant Biol. 2024;24:1100.

62. Naes T, Tomic O, Mevik B-H, Martens H. Path modelling by sequential PLS regression. J Chemometrics. 2011;25:28–40.

63. Liland KH, Næs T, Indahl UG. ROSA—a fast extension of partial least squares regression for multiblock data analysis. Journal of Chemometrics. 2016;30:651–62.

64. Kandpal LM, Mouazen AM, Masithoh RE, Mishra P, Lohumi S, Cho B-K, et al. Sequential data-fusion of near-infrared and mid-infrared spectroscopy data for improved prediction of quality traits in tuber flours. Infrared Physics & Technology. 2022;127:104371.

65. Kang C, Diverres G, Karkee M, Zhang Q, Keller M. Assessing grapevine water status through fusion of hyperspectral imaging and 3D point clouds. Computers and Electronics in Agriculture. 2024;226:109488.

66. Ji F, Li F, Hao D, Shiklomanov AN, Yang X, Townsend PA, et al. Unveiling the transferability of PLSR models for leaf trait estimation: lessons from a comprehensive analysis with a novel global dataset. New Phytologist. 2024;243:111–31.

67. Grzybowski M, Wijewardane NK, Atefi A, Ge Y, Schnable JC. Hyperspectral reflectance-based phenotyping for quantitative genetics in crops: Progress and challenges. Plant Communications. 2021;2:100209.

68. Jablonszky M, Garamszegi LZ. The effect of repeated measurements and within-individual variance on the estimation of heritability: a simulation study. Behav Ecol Sociobiol. 2024;78:18.

69. Zhang H, Wang L, Jin X, Bian L, Ge Y. High-throughput phenotyping of plant leaf morphological, physiological, and biochemical traits on multiple scales using optical sensing. The Crop Journal. 2023;11:1303–18.

70. Wen T, Li J-H, Wang Q, Gao Y-Y, Hao G-F, Song B-A. Thermal imaging: The digital eye facilitates high-throughput phenotyping traits of plant growth and stress responses. Science of The Total Environment. 2023;899:165626.

71. Tattaris M, Reynolds MP, Chapman SC. A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. Front Plant Sci. 2016;7.

72. Coupel-Ledru A, Pallas B, Delalande M, Segura V, Guitton B, Muranty H, et al. Tree architecture, light interception and water-use related traits are controlled by different genomic regions in an apple tree core collection. New Phytologist. 2022;234:209–26.