Leaf hyperspectral data and different regression models to estimate photosynthetic parameters in two leguminous crops

Ma. Luisa Buchaillot, David Soba, Tianchu Shu, Juan Liu, Iker Aranjuelo, José Luis Araus, Shawn C. Kefauver, Álvaro Sanz-Saez

UNIVERSITAT DE BARCELONA

AUBURN UNIVERSITY

Objectives

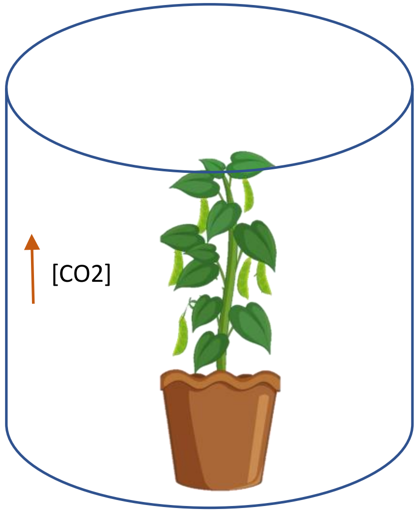
(1) Predict photosynthetic capacity parameters, such as midday photosynthesis, leaf chlorophyll content (LCC), normalized midday photosynthesis with LCC, Vmax, and Jmax of two species of legumes using leaf reflectance spectra (VIS-NIR-SWIR) with advanced statistical models including Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso).

(2) Simulate if equipment, such as UAV hyperspectral cameras or lower cost field spectrometers with lesser spectrum coverage limited to VIS-NIR, NIR-SWIR or SWIR, and simulated Sentinel-2 bands would be able to detect with similar accuracy the same parameters estimated with the ASD FieldSpec leaf reflectance.

Materials and Methods

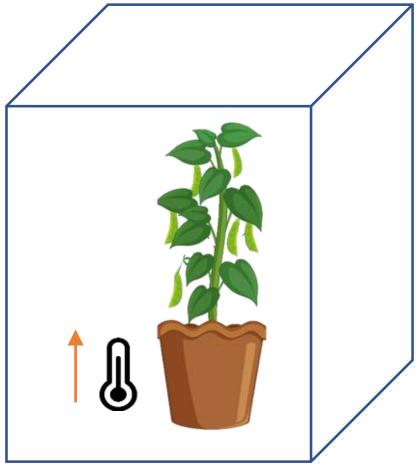
Third Experiment:

First Experiment:

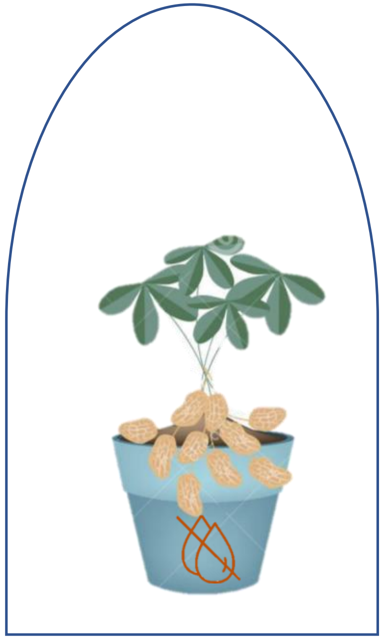


2 varieties

Second Experiment:



4 varieties

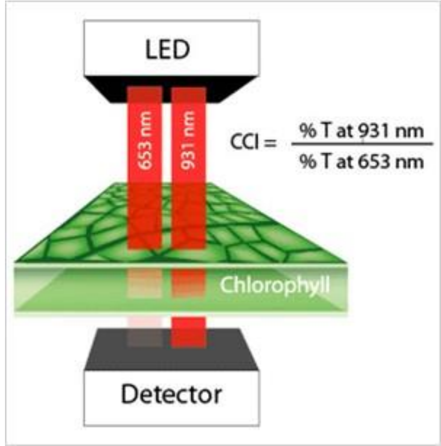


6 varieties

Crop	Peanut	Soybean ↑ [CO_2]	Soybean ↑
Sowing day	April 21st 2019	May 16th 2019	May 1st 2019
Harvest day	August 28th 2019	R5 August 1st 2019	August 19th 2019
Varieties (amount and name)	6 varieties AU16-28 (Tol), AU17 (Tol), 18H19-3738 (Tol), G-06-G (Sen), AU8-19 (Sen), AU18-21 (Sen).	2 varieties PI398223 (high WUE) and PI (PI567201A). You have 4 blocks but in each block we took 2 plants.	4 varieties PI360846 (high Temp Tolerant), DS25-1 (Tolerant), Pi458098 (High Temp sensitive), AG48x9 (commercial) in 4 replicates
Treatment	Well water, 80% of soil water content measured by gravimetric methods. Water stress 30% of SWC.	Ambient (410 umol mol-1 CO_2) and Elevated (610 umol mol-1 CO_2)	High night temperature for two weeks after full Bloom. Control Chamber 30/20C day nigh. Treatment Chamber 30/30C day night. Light intensity in the chamber 1000umol PAR. HR 70/60% day night
Type of greenhouse	Glass greenhouse	Open Top Chamber	Conviron Growth Chamber
Photosynthetic Parameters (Li-COR and FieldSpect4)	• July 12th-13th 2019 • July 26th-27th 2019	• July 15th-16th 2019 • July 26th-27th 2019	• June 15th 2019
Relative chlorophyll content (SPAD)	• July 12th-13th 2019 • July 26th-27th 2019	• July 16th 2019 • July 26th 2019	• June 15th 2019

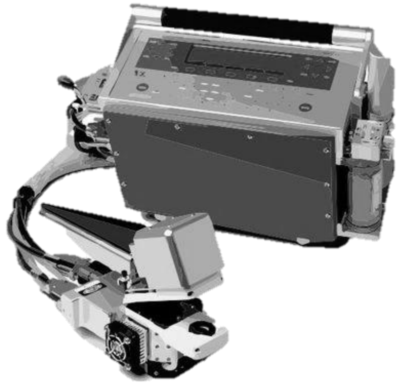
Materials and Methods

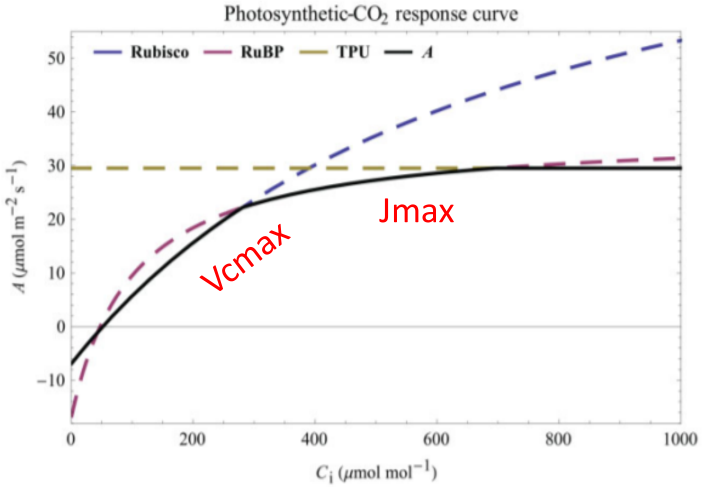
SPAD-502 chlorophyll meter



1 min.

Licor-6400 for Leaf gas exchange

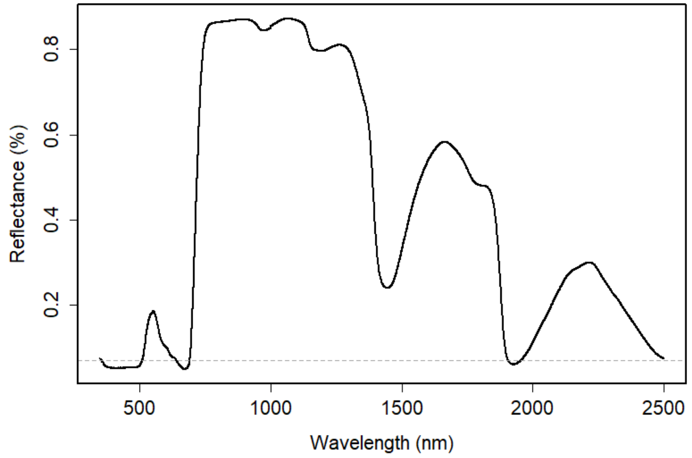




40 min.

FieldSpect4 full range spectrometer

Leaf reflectance spectrum



1 min.

Materials and Methods

Machine Learning models

DIMENSION REDUCTION METHOD

- Partial Least Squares Regression (PLSR)

SHRINKAGE METHOD

- Least Absolute Shrinkage and Selection Operator (Lasso)

HIGH-DIMENSIONAL METHODS

- Bayesian Ridge (BR)
- Automatic Relevance Determination Regression (ARDR)

Results

Table 1. Minimum and maximum of the Photosynthesis (μmol $m^{-2}s^{-1}$), Leaf Chlorophyll Content (LCC) (arbitrary unit), Photosynthesis/LCC (μmol $m^{-2}s^{-1}spad^{-1}$), Vcmax (μmol $m^{-2}s^{-1}$) and Jmax (μmol $m^{-2}s^{-1}$) per each treatment. **A)** Experiment 1: 2 varieties of soybean grown under 410 ppm and 610 ppm of $[CO_2]$. **B)** Experiment 2, 4 soybean varieties grown under control (20°C) and high (30°C) night temperature. **C)** Six varieties of peanut grown under Well Water (WW, 80% SWC) and Water Stress (WS, 30% SWC). Significant differences through ANOVA.

A)

Varieties	**Treatment**	**Midday Photosynthesis (μmol $m^{-2}s^{-1}$)**		**LCC (arbitrary unit)**		**Midday Photosynthesis/LCC (μmol $m^{-2}s^{-1}spad^{-1}$)**		**Vcmax (μmol $m^{-2}s^{-1}$)**		**Jmax (μmol $m^{-2}s^{-1}$)**	
amount	**Soybean**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**
2	410	17	31.3	34.55	49.11	0.348	0.942	190.5	348.4	174.7	263.7
2	610	21.5	36.2	40.7	51.35	0.475	0.806	182.9	322.7	185.3	251.4
ANOVA	**[CO^2] (p-value)**	**0.001**		**0.015**		**0.010**		*0.242*		*0.569*	
ANOVA	**Varieties (p-value)**	*0.344*		*0.130*		*0.578*		**0.092**		*0.173*	
ANOVA	**[CO^2]*Varieties (p-value)**	*0.678*		**0.077**		*0.513*		*0.733*		*0.646*	

B)

Varieties	**Treatment**	**Midday Photosynthesis (μmol $m^{-2}s^{-1}$)**		**LCC (arbitrary unit)**		**Midday Photosynthesis/LCC (μmol $m^{-2}s^{-1}spad^{-1}$)**		**Vcmax (μmol $m^{-2}s^{-1}$)**		**Jmax (μmol $m^{-2}s^{-1}$)**	
amount	**Soybean**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**
4	Low T°	13.3	30.72	34	53.1	0.292	0.665	48	124	78	165
4	High T°	11.5	32.68	37.2	53.9	0.287	0.661	63	135	61	165
ANOVA	**Temperature (p-value)**	*0.103*		*0.522*		**0.031**		*0.833*		*0.624*	
ANOVA	**Varieties (p-value)**	**0.010**		**0.002**		**0.031**		*0.303*		**0.042**	
ANOVA	**Temp*Varieties (p-value)**	*NS*		*NS*		*NS*		*NS*		*NS*	

C)

Varieties	**Treatment**	**Midday Photosynthesis (μmol $m^{-2}s^{-1}$)**		**LCC (arbitrary unit)**		**Midday Photosynthesis/LCC (μmol $m^{-2}s^{-1}spad^{-1}$)**		**Vcmax (μmol $m^{-2}s^{-1}$)**		**Jmax (μmol $m^{-2}s^{-1}$)**	
amount	**Peanut**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**	**min**	**max**
6	WW	11.4	26.4	42.3	52.45	0.242	0.516	84.27	171.3	83.05	206.1
6	WS	5.05	19.4	46.36	51.17	0.092	0.355	64.38	162.7	79.39	188.8
ANOVA	**Drought (p-value)**	**0.000**		**0.001**		**0.001**		*0.460*		**0.000**	
ANOVA	**Varieties (p-value)**	*0.154*		**0.001**		*0.478*		*0.196*		**0.092**	
ANOVA	**Drought*Varieties (p-value)**	*0.884*		*0.353*		*0.740*		**0.094**		*0.352*	

Results

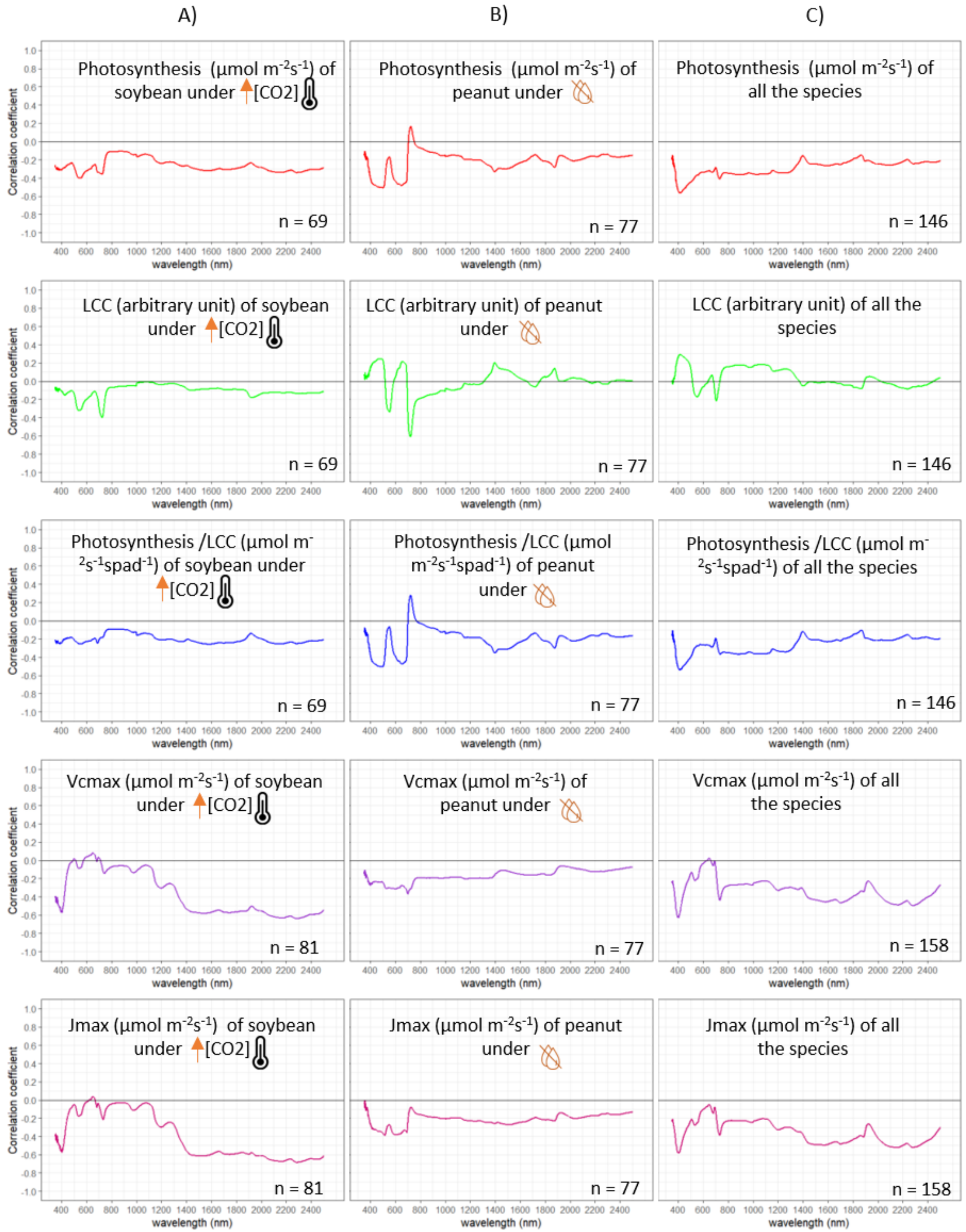


Figure 1. Pearson correlation coefficients between the photosynthetic parameters and each wavelength from the leaf reflectance spectrum for each species and all together. **A)** Soybean varieties under two treatments, one under high [CO2] and the other under high temperature. **B)** Peanut varieties under water stress. **C)** Soybean and peanut varieties together. Each graphic present in the x axis the wavelength spectrum and in the y axis the Pearson correlation coefficient from -1 to 1.

Results

Table 2. Coefficient of determination (R^2) and Root Mean Squared Error (RMSE) of the photosynthesis (µmol $m^{-2}s^{-1}$), Leaf chlorophyll content (LCC) and photosynthesis/LCC (µmol $m^{-2}s^{-1}spad^{-1}$) Vcmax (µmol $m^{-2}s^{-1}$) and Jmax (µmol $m^{-2}s^{-1}$) of all the species together based on leaf reflectance spectra VIS-NIR-SWIR (350-2500nm) through advance regression models : Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso).

Model	Estimation of Photosynthesis (µmol $m^{-2}s^{-1}$) of all species		Estimation of LCC (arbitrary unit) of all the species		Estimation of Photosynthesis /LCC (µmol $m^{-2}s^{-1}spad^{-1}$) of all species		Estimation of maximum rate of carboxylation of RuBP (Vcmax) of all species		Estimation of maximum rate of electron transport driving RuBP regeneration (Jmax) of all species	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
PLSR	0.51	5.41	0.49	4.14	0.27	0.14	0.7	42.79	0.55	30.59
BR	0.4	5.98	0.08	5.56	0.37	0.13	0.59	50.15	0.51	31.72
ARDR	0.29	6.53	0.34	4.7	0.05	0.16	0.56	52.03	0.52	31.62
LASSO	0.34	6.29	0.17	5.26	--	--	0.59	50.11	0.42	34.74

Results

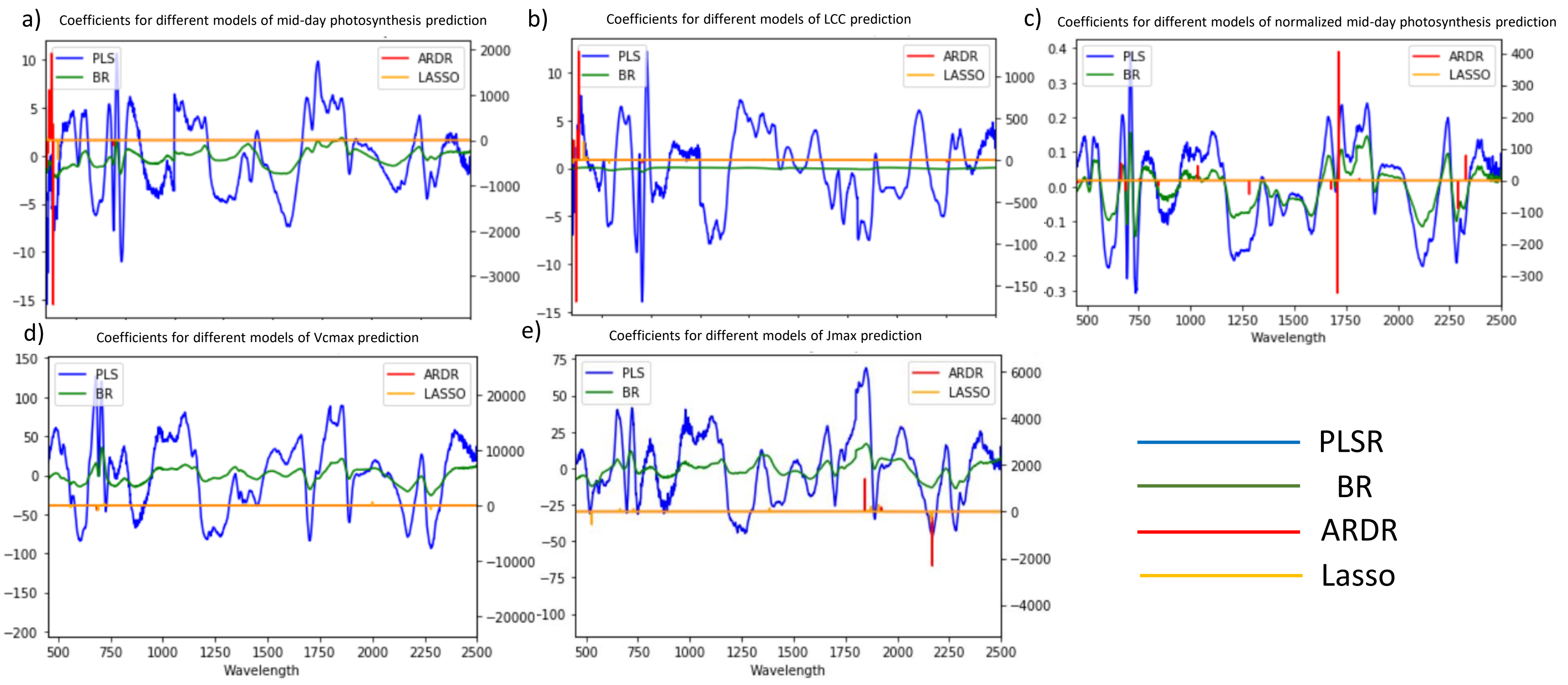


Figure 2. PLSR (blue), BR (green), ARDR (red) and Lasso (yellow) models build generated spectral-specific coefficients for **a)** mid-day photosynthesis (μmol $m^{-2}s^{-1}$), **b)** Leaf chlorophyll content (LCC) and **c)** normalized mid-day photosynthesis (μmol $m^{-2}s^{-1}spad^{-1}$) **d)** maximum rate of carboxylation of RuBP (μmol $m^{-2}s^{-1}$) (Vc_{max}) and **e)** maximum rate of electron transport driving RuBP regeneration (μmol $m^{-2}s^{-1}$) (J_{max}) of all the species together.

Results

Table 3. Coefficient of determination (R^2) and Root Mean Squared Error (RMSE) of the maximum rate of carboxylation of RuBP (µmol $m^{-2}s^{-1}$) (Vc_{max}) and maximum rate of electron transport driving RuBP regeneration (µmol $m^{-2}s^{-1}$) (J_{max}) maximum of the first and second experiment (soybean) together based on leaf reflectance spectra at different ranges: VIS-NIR (350-1000nm), NIR-SWIR (1000-2500nm), SWIR (1400-2500nm) and Sentinel-2 bands through advance regression models : Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso).

Estimation of maximum rate of carboxylation of RuBP (Vcmax) of soybean								
n = 81	**Vc_{max}** from 350 to 1000 nm		**Vc_{max}** from 1000 to 2500 nm		**Vc_{max}** from 1400 to 2500 nm		Simulation of Sentinel-2	
Model	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
PLSR	0.49	60.23	--	--	0.37	69.33	0.45	62.24
BR	0.63	46.57	0.29	70.59	0.39	65.46	0.6	53.19
ARDR	0.54	56.85	0.47	61.16	0.48	60.4	0.39	65.5
LASSO	0.64	50.57	0.46	61.46	0.45	62.45	0.66	49.01
Estimation of maximum rate of electron transport driving RuBP regeneration (Jmax) of soybean								
n = 81	**J_{max}** from 350 to 1000 nm		**J_{max}** from 1000 to 2500 nm		**J_{max}** from 1400 to 2500 nm		Simulation of Sentinel-2	
Model	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
PLSR	0.42	41.03	0.35	43.39	0.4	41.64	0.32	44.25
BR	0.65	31.92	0.55	36.07	0.55	36.06	0.54	36.42
ARDR	0.58	34.93	0.56	35.5	0.61	33.47	0.51	37.69
LASSO	0.56	35.73	0.55	36.13	0.56	35.88	0.59	34.24

Results

Table 4. Coefficient of determination (R^2) and Root Mean Squared Error (RMSE) of maximum rate of carboxylation of RuBP (µmol $m^{-2}s^{-1}$) (Vc_{max}) and maximum rate of electron transport driving RuBP regeneration (µmol $m^{-2}s^{-1}$) (J_{max}) of the all species together based on leaf reflectance spectra at different ranges: VIS-NIR (350-1000nm), NIR-SWIR (1000-2500nm), SWIR (1400-2500nm) and Sentinel-2 bands through advance regression models : Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso).

Estimation of maximum rate of carboxylation of RuBP (Vcmax) of all the species								
n = 158	**Vc_{max}** from 350 to 1000 nm		**Vc_{max}** from 1000 to 2500 nm		**Vc_{max}** from 1400 to 2500 nm		Simulation of Sentinel-2	
Model	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
PLSR	0.56	51.94	0.65	46.49	0.61	49.23	0.57	51.62
BR	0.6	49.5	0.62	48.1	0.6	49.35	0.31	65.24
ARDR	0.52	54.03	0.55	59.49	0.58	50.67	0.31	65.65
LASSO	0.56	52.17	0.53	53.71	0.53	53.63	0.5	55.61
Estimation of maximum rate of electron transport driving RuBP regeneration (Jmax) of all the species								
n = 158	**J_{max}** from 350 to 1000 nm		**J_{max}** from 1000 to 2500 nm		**J_{max}** from 1400 to 2500 nm		Simulation of Sentinel-2	
Model	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
PLSR	0.32	41.74	0.47	36.89	0.51	35.57	0.41	38.9
BR	0.35	39.73	0.49	36.01	0.47	36.77	0.41	38.71
ARDR	0.11	47.69	0.52	35.1	0.51	35.36	0.39	39.43
LASSO	0.42	38.41	0.41	38.89	0.36	40.56	0.41	38.38

Conclusions

- Regarding with the first aim: The results suggest that we can predict the Vcmax and Jmax with a coefficient of determination greater than 50% through the four machine learning models Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso) based on the full range leaf reflectance spectrum (350-2500nm) of the two legumes.

- Concerning the second aim: We tested different sensors with lower ranges of wavelengths such as UAV level hyperspectral cameras that have shorter spectrums like VIS-NIR, NIR-SWIR or SWIR, and simulation of Sentinel-2 bands to estimate the Vcmax and Jmax with the four advance models. The results suggest that the models vary depending on the treatments (high CO_2, high temperature, and drought) that we did in the different experiments; different sensors work better than others with different advanced models.

- In future work, it may be useful to continue working across different species under different treatments and complete a much larger data base with all Vcmax and Jmax at different sites, as this information could, in the future, be used could predict these parameters with an estimation accuracy of around 80%, and reduce the time for 40 min (Li-COR 6-400) to 1 min (spectral or hyperspectral imaging) per plant.

Supplemental

Supplemental Table 1. Coefficient of determination (R^2) and Root Mean Squared Error (RMSE) of the photosynthesis (µmol $m^{-2}s^{-1}$), Leaf chlorophyll content (LCC) and photosynthesis/LCC (µmol $m^{-2}s^{-1}spad^{-1}$) Vcmax (µmol $m^{-2}s^{-1}$) and Jmax (µmol $m^{-2}s^{-1}$) of peanut and soybean together based on leaf reflectance spectra VIS-NIR-SWIR (350-2500nm) through advance regression models : Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR) and Least Absolute Shrinkage and Selection Operator (Lasso).

Model	Estimation of Photosynthesis (µmol $m^{-2}s^{-1}$) of peanut		Estimation of LCC (arbitrary unit) of the peanut		Estimation of Photosynthesis /LCC (µmol $m^{-2}s^{-1}spad^{-1}$) of peanut		Estimation of maximum rate of carboxylation of RuBP (Vcmax) of peanut		Estimation of maximum rate of electron transport driving RuBP regeneration (Jmax) of peanut	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
PLSR	--	--	0.46	3.77	--	--	--	--	--	--
BR	--	--	0.47	3.72	--	--	--	--	--	--
ARDR	--	--	0.48	3.69	--	--	--	--	--	--
LASSO	--	--	0.42	3.91	--	--	--	--	--	--

Model	Estimation of Photosynthesis (µmol $m^{-2}s^{-1}$) of soybean		Estimation of LCC (arbitrary unit) of soybean		Estimation of Photosynthesis /LCC (µmol $m^{-2}s^{-1}spad^{-1}$) of soybean		Estimation of maximum rate of carboxylation of RuBP (Vcmax) of soybean		Estimation of maximum rate of electron transport driving RuBP regeneration (Jmax) of soybean	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
PLSR	0.05	5.9	--	--	--	--	0.63	51.04	0.49	38.25
BR	0.21	5.38	--	--	--	--	0.64	50.16	0.61	33.61
ARDR	0.22	5.35	--	--	--	--	0.49	60.03	0.59	34.64
LASSO	0.22	5.36	--	--	--	--	0.63	51.1	0.6	33.83