DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Title: Mechanistic links between physiology and spectral reflectance enable pre-visual detection of oak wilt and drought stress

Supplementary materials

Appendix S1. Expanded materials and methods.

S1.1. Map of Cedar Creek showing the location of all oak wilt infected (purple) and Control (green) seedlings.

S1.2 Expanded methods for measurements of water relations and active xylem staining.

Water relations

Immediately after measuring leaves' water potential, we weighed them to obtain fresh weight. We then momentarily submerged them in a deionized water reservoir on a scale to obtain their volume using the water displacement technique (Sapes et al., 2019). Leaves were then blotted dry with paper towels, sealed inside a Ziploc bag, and placed in the fridge with their petioles submerged in water-filled vials for 8 hours to allow full recovery of turgor (Boyer et al., 2008). We measured their saturated weight, then oven-dried them at 70 ºC until they reached constant mass, to obtain their dry weight. To prevent possible undersaturation artifacts due to cellular damage and embolism, we measured RWC as the ratio between current water fraction (WF) (calculated as (Fresh weight - Dry weight)/Dry weight) and the saturated water fraction ($SWF_{initial}$) (calculated as (Saturated weight - Dry weight)/Dry weight) before applying treatments, expressed in percentage units. VWC was calculated following Sapes (2019) as ((Fresh weight - Dry weight)/Fresh volume)*100. We calculated LRC for each individual as 100*(1- ($SWF_{current}$/ $SWF_{initial}$)). LRC values above 0 indicate the degree of lost capacity to recover full turgor. In the potted experiment, we obtained the $SWF_{initial}$ of each tree before trees experienced any stress (i.e., before treatment). At Cedar Creek, we did not know the stress history of trees, which prevented us from ensuring that $SWF_{initial}$ were truly saturated. Instead, we used the mean saturated water fraction values of the control trees at the start of the experiment as a proxy for individual $SWF_{initial}$ (population-level approach). When using a population-level approach, some LRC values can be slightly negative because some plants will inevitably have lower $SWF_{initial}$ than the population mean. However, we could still observe clear patterns between treatments despite the added noise (Appendix S2).

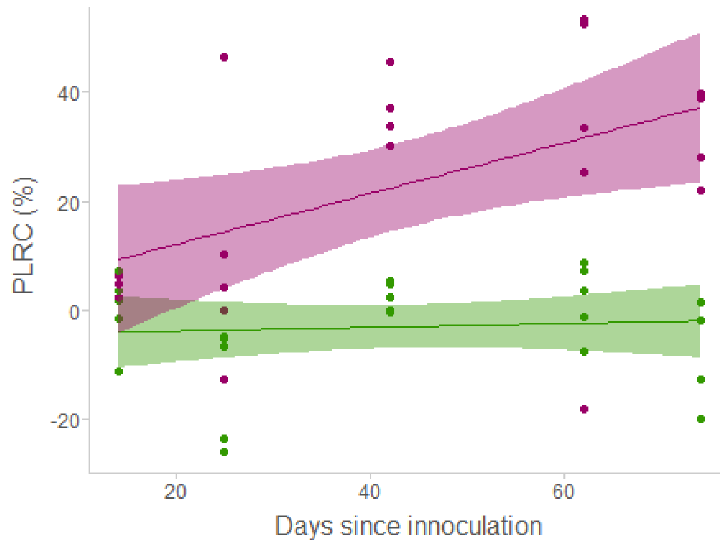
For the additional leaves used to measure electrolyte leakage (EL), two hole-punched sections per leaf were obtained and placed in deionized water-filled 15 mL vials. Tubes were shaken at 110 rpm for 20 hours at 5 C to limit metabolic activity. Electrical conductivity (C1) was measured using a PC 510 conductivity probe. Then, samples were autoclaved for 30 min to kill all living cells, shaken for an additional 20 hours, and measured again (C2). Lastly, EL was calculated as 100*(C1/C2).

Active xylem staining

Stem segments were cut and re-cut under degassed water to prevent embolism formation and then connected to a reservoir holding degassed water. Stem segments were then loaded into vials containing 0.1% safranin solution and the water reservoir was lowered to generate a pulling force of 2 kPa which infiltrated functional conduits with the stain. Once the stain solution reached the other end of the stem, excess dye inside the conduits was pushed out by raising the reservoir to reverse the pressure gradient. We then cut a cross-section of the stems to measure the amount of stained (functional) and non-stained (dysfunctional) xylem. To remove emboli, stems were flushed with degassed water at ca. 100 kPa using a syringe mounted in a hand-held caulking gun until bubbles stopped coming out of the ends. The staining process was repeated to measure the area of xylem that was embolized by comparing cross sections before and after flushing. Conduits filled with tyloses due to oak wilt remained unstained. We used ImageJ (62) to measure the percent functional xylem area before flushing and the percent clogged xylem area after flushing. Additionally, we divided each cross-section into 16 sections and counted the number of contiguous obstructed xylem sections to quantify the degree of localized versus multiple xylem obstruction patterning.

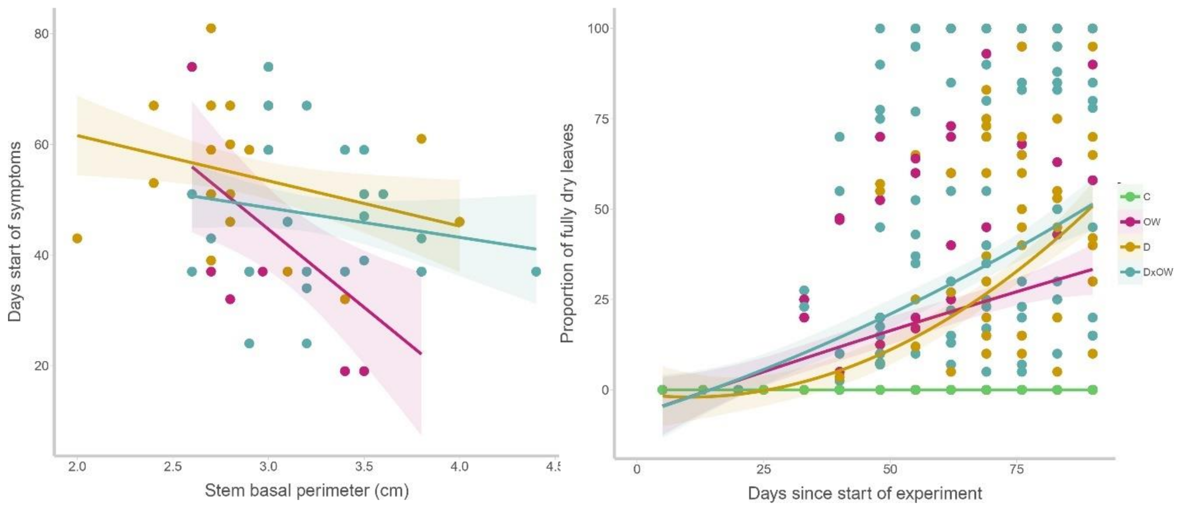
51 **Appendix S2. Patterns of population-level loss of rehydration capacity.** When using a population-level approach, some LRC values can be slightly negative because some plants will inevitably have lower SWF $_{initial}$ than the population mean. However, clear patterns can still be observed between oak wilt (purple) and control (green) treatments despite the added noise.

55



DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Appendix S3. Trees with larger stems developed symptoms earlier in time which led to large within-treatment variability in canopy decline. Panel A: Relationship between the number of days until the start of visual symptom appearance and the stem basal perimeter. Oak wilt infected trees showed the largest reduction in time to symptom appearance as a function of plant size. Panel B: Relationship between the proportion of fully dry leaves and time since the start of the experiment. The large within-treatment variability in proportion of dry leaves is explained by the variability in plant size observed in panel A. Each point represents a tree.



Model and Factors	Model Type	Estimate	2.50%	97.50%	p-value	d.f. (res.)	Adjusted R square
Days until start of symptoms ~ stem basal perimeter x treatment	LM				<0.001	137	0.19
Intercept		129.3109	85.98294	172.6389	<0.001		
stem basal perimeter		-28.2213	-42.0578	-14.3848	<0.001		
treatmentD		-51.3685	-100.206	-2.53131	0.039		
treatmentDxOW		-64.76	-115.363	-14.1567	0.013		
stem basal perimeter x treatmentD		20.04133	4.23506	35.84761	0.013		
stem basal perimeter x treatmentDxOW		22.87994	6.86211	38.89777	0.005		

Appendix S4. Custom mathematical function developed to select a reduced set of wavelengths while still representing the shape of a plant reflectance spectrum across the 400-2400 nm range.

Objective:

Generate a function that selects the minimum number of wavelengths necessary to describe a whole plant spectrum profile.

Logic statement 1: Reflectance values at neighboring wavelengths are highly correlated, thus reflectance at a given wavelength is similar to that of the previous one.

Logic statement 2: The shape of a spectrum profile is determined by the change in reflectance from one wavelength to the next.

Logic statement 3: The amount of new information about the shape of the spectrum provided by a target wavelength is proportional to the amount of change in reflectance between contiguous wavelengths.

Based on logic statements 2 & 3: if reflectance does not change from one wavelength to the next, the shape remains constant. Thus, the next wavelength does not provide new information on the shape and can be skipped.

Conclusion: the change in reflectance between two contiguous wavelengths can be used to calculate how many wavelengths can be skipped at any region of the spectra without missing critical information about the shape of the spectra.

Constraints imposed by user:

Constraint 1: Selected wavelengths should range from 400 to 2400 nm.

Constraint 2: Wavelengths within 5 nm from two target wavelengths can be skipped due to high correlation.

Constraint 3: Distance between two selected wavelengths should not be greater than 40 nm to avoid missing narrow but important features of the spectrum.

Constraint 4: User must be able to control the magnitude of the jumping distance that determines which wavelengths to skip.

Variables resulting from logic statements and constraints:

$jumping\ distance = amount\ of\ nanometers\ to\ skip\ between\ two\ selected\ wavelengths$

$min = minimum\ distance\ between\ two\ selected\ wavelengths$

$max = maximum\ distance\ between\ two\ selected\ wavelengths$

$R_x = Reflectance\ at\ wavelength\ X$

$R_{x+1} = Reflectance\ at\ wavelength\ X + 1$

$a = tunning\ parameter\ 1$

$b = tunning\ parameter\ 2$

Function:

$$jumping\ distance = min + \left(\left(\log\left(\frac{a}{1 + abs(R_{x+1} - R_x) * b}\right)\right) * (max/min)\right)$$

Where *min* and *max* were set to 5 and 40 nm, respectively. *a* and *b* were set to 13 and 10,000 respectively based on our range of reflectance change. Reflectance change ranged from 0.000000070 to 0.001065945.

Note that as change in reflectance increases, the denominator of the logarithm approaches values close to *a* and yields a logarithm value close to 0. Thus, jumping distance is equal to *min*.

The blue side of the equation therefore determines the jumping distance during high changes in reflectance based on the *min* parameter. The red side of the equation modulates the distance of the jump at intermediate and low values of change and provides user control through the tunning parameters *a* and *b*. The yellow side of the equation regulates the size of the jumps based on the *max* parameter.

The operators used within the function have no particular meaning other than to achieve the desired flexible behavior: determine the jumping distance needed to select a small subset of wavelengths across the spectral profile while retaining the shape and features of the whole spectrum.

This formula was applied to the reflectance of control trees as they best represented the typical spectra profile of plants. We iterated through different tunning until we found a set of parameters that returned jumping distances that yielded a subset of wavelengths representative of the whole spectrum. The result was 94 selected wavelengths with jumping distances ranging from 6 to 26 nm across the spectrum.

DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Appendix S5. Generalized (GLM) and linear (LM) models showing visual and physiological decline as a function of days since the start of visual symptom appearance. Plots are shown in figure 2.

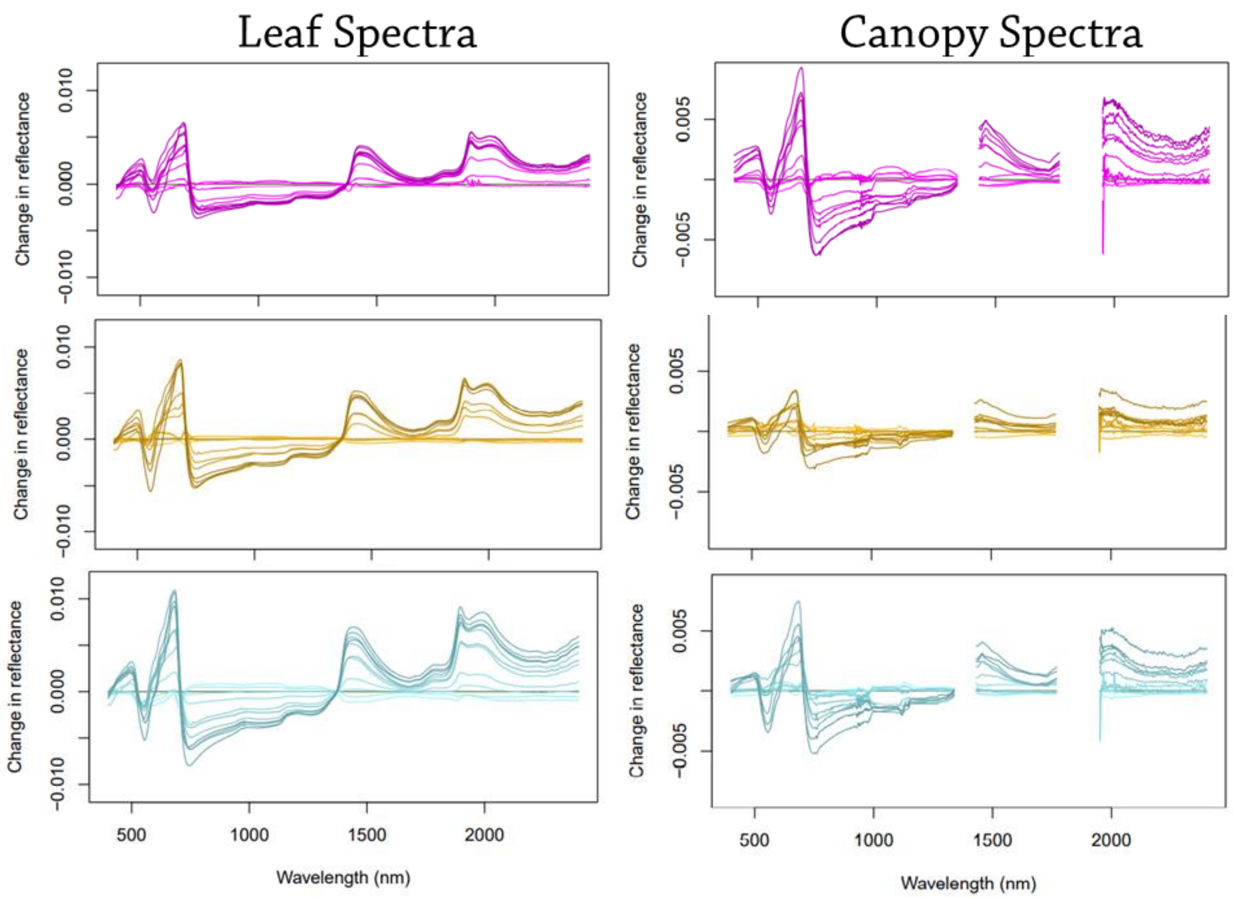
			95% C.I. Estimates				
Model and Factors	**Model Type**	**Estimate**	**2.50%**	**97.50%**	***p-value***	**d.f. (res.)**	**AIC**
Proportion of healthy leaves/100 ~ Days since start symptoms	GLM					655	235.81
Intercept		1.2013	0.87225	1.56117	**<0.001**		
Days since start symptoms		-0.14877	-0.17544	-0.12575	**<0.001**		
Percent electrolyte leakage/100 ~ Days since start symptoms	GLM					141	142.8
Intercept		-0.75257	-1.1217	-0.40039	**<0.001**		
Days since start symptoms		0.01406	0.00232	0.02644	**0.022**		
Percent loss rehydration capacity/100 ~ Days since start symptoms	GLM					137	121.85
Intercept		-0.4926	-0.88104	-0.12007	**0.011**		
Days since start symptoms		0.03197	0.01873	0.04674	**<0.001**		
Relative water content/100 ~ Days since start symptoms	GLM					138	121.98
Intercept		-0.01688	-0.38909	0.35381	0.929		
Days since start symptoms		-0.03433	-0.04918	-0.02107	**<0.001**		
Volumetric water content/100 ~ Days since start symptoms	GLM					139	145.18
Intercept		-0.6392	-1.01331	-0.28247	**<0.001**		
Days since start symptoms		-0.01974	-0.03266	-0.00774	**0.002**		

DOI of published paper: https://doi.org/10.1073/pnas.2316164121

			95% C.I. Estimates				
Model and Factors	**Model Type**	**Estimate**	**2.50%**	**97.50%**	***p-value***	**d.f. (res.)**	**Adjusted R square**
Water potential ~ poly(Days since start symptoms, 3)	LM				**<0.001**	138	0.61
Intercept		-4.50519	-4.87256	-4.13782	**<0.001**		
Days since start symptoms		-70.792	-82.0639	-59.5201	**<0.001**		
Days since start symptoms^2		-0.98381	-13.7986	11.83097	0.88		
Days since start symptoms^3		19.97869	8.03137	31.92601	**0.001**		
Fv/Fm ~ days_since_start_symptoms	LM				**<0.001**	115	0.35
Intercept		0.6606	0.63628	0.68491	**<0.001**		
Days since start symptoms		-0.00296	-0.0037	-0.00222	**<0.001**		

DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Appendix S6. Changes in leaf and canopy spectral reflectance over time in oak wilt (purple), drought (yellow), and drought x oak wilt (blue) treatments. Values represent the difference in reflectance between the average spectra of a treatment and the average spectra of controls. Positive values indicate higher reflectance in treatment than controls while negative values indicate lower reflectance. Darker color tones within a plot represent later measurements in time.



DOI of published paper: https://doi.org/10.1073/pnas.2316164121

136 **Appendix S7. Performance metrics for the partial least square regression (PLSR) models.** Variables described correspond to relative water content (RWC), loss of rehydration capacity (LRC), volumetric water content (VWC), electrolyte leakage (EL), maximum efficiency of photosystem II (FvFm), and water potential (WP). Metrics described correspond to medians of the number of components used to build the models, maximum and minimum measured values of the response variable, coefficient of determination (R2), p-value, root mean square error (RMSE) of predicted values and root mean square error percentage (RMSEP) of predicted values, slope of predicted vs measured values, and bias across model iterations. Metrics are shown for both, testing using data from the outdoor potted experiment and independent validation using data from Cedar Creek.

143 **All trees**

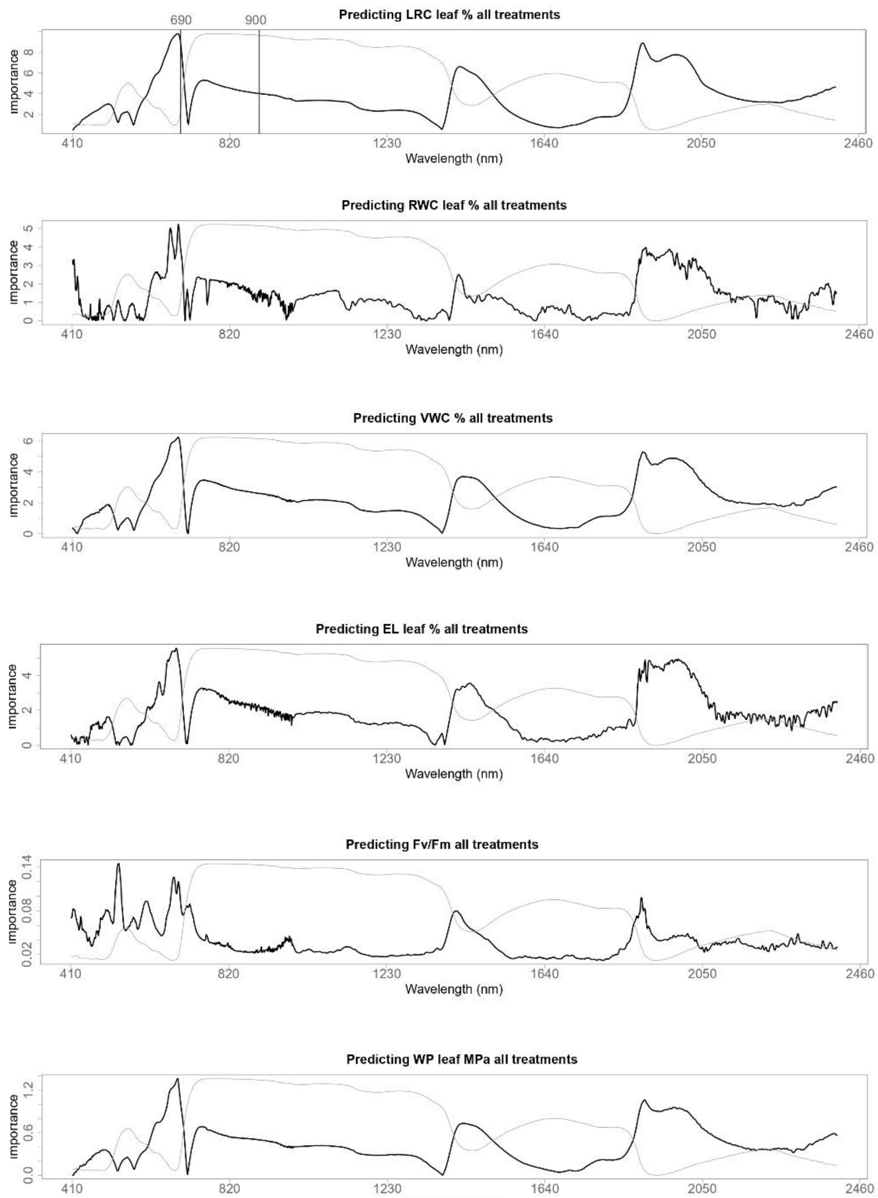
				Testing						Independent Validation				
Variable	**components**	**max**	**min**	**R2**	**P-value**	**RMSE**	**RMSEP**	**slope**	**bias**	**R2**	**P-value**	**RMSE**	**RMSEP**	**bias**
RWC	4	88.72	1.76	0.87	<0.001	9.95	11.47	0.99	0.00	0.50	<0.001	26.99	31.17	19.09
LRC	4	91.84	-0.98	0.84	<0.001	11.07	11.98	0.99	-0.01	-	-	-	-	-
VWC	2	71.73	2.33	0.81	<0.001	7.40	10.76	0.99	-0.01	0.48	<0.001	13.72	25.32	-5.32
EL	3	80.57	14.87	0.75	<0.001	8.50	13.09	1.01	0.01	-	-	-	-	-
FvFm	4	0.86	0.07	0.54	<0.001	0.13	16.16	0.97	0.00	0.47	<0.001	0.17	22.81	-0.09
WP	2	-0.25	-8.00	0.33	<0.001	1.53	19.77	0.89	0.00	0.39	<0.001	2.68	36.67	-1.56

144 **Without control trees**

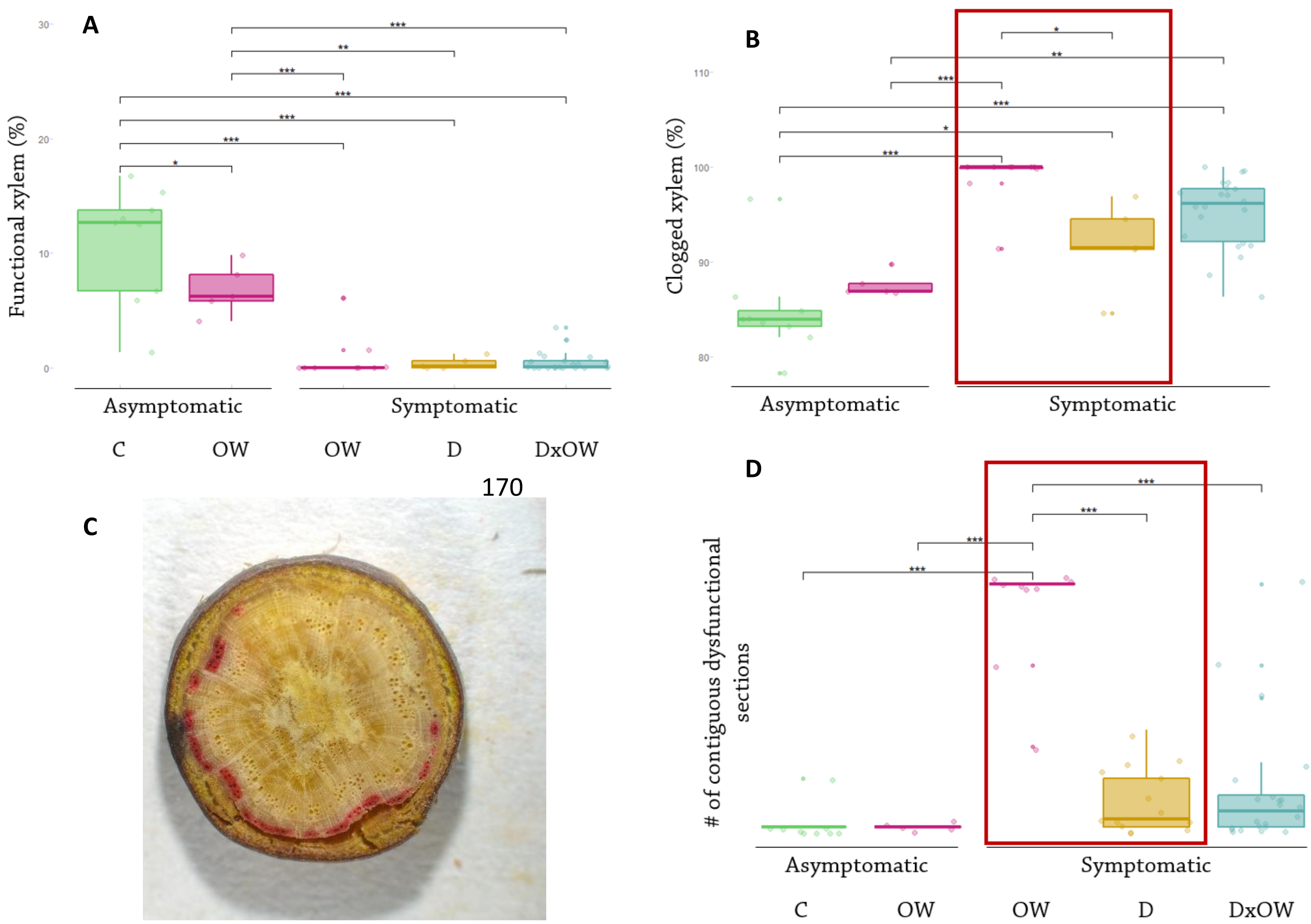
				Testing						Independent Validation				
Variable	**components**	**max**	**min**	**R2**	**P-value**	**RMSE**	**RMSEP**	**slope**	**bias**	**R2**	**P-value**	**RMSE**	**RMSEP**	**bias**
RWC	4	88.72	1.76	0.88	<0.001	9.74	11.30	1.00	-0.01	0.50	<0.001	26.39	30.48	17.83
LRC	3	91.84	-0.98	0.84	<0.001	11.25	12.29	1.00	0.09	-	-	-	-	-
VWC	2	71.73	2.33	0.80	<0.001	7.80	11.38	1.01	-0.03	0.48	<0.001	13.34	24.62	-4.45
EL	3	80.57	14.87	0.75	<0.001	8.84	13.61	0.98	-0.03	-	-	-	-	-
FvFm	3	0.86	0.07	0.49	<0.001	0.13	17.14	0.92	0.00	0.49	<0.001	0.17	22.32	-0.09
WP	2	-0.40	-8.00	0.32	<0.001	1.58	20.79	0.98	0.00	0.40	<0.001	2.63	36.02	-1.57

DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Appendix S8. Importance of wavelengths on predicting physiological processes related to oak wilt and drought. Variables described correspond to relative water content (RWC), loss of rehydration capacity (LRC), volumetric water content (VWC), electrolyte leakage (EL), maximum efficiency of photosystem II (FvFm), and water potential (WP). Black line indicates variable importance of predictors (VIP). Gray line shows a representative leaf spectrum from a control plant for reference. Vertical lines indicate wavelengths 690 and 900 commonly used in the Renormalized Difference Vegetation Index (RDVI).



Appendix S9. Vascular embolism due to drought and occlusion due to oak wilt. Panel A: Trees infected with oak wilt that showed no visual symptoms (asymptomatic) presented some functional xylem -but less than controls- whereas symptomatic trees lost all their functional xylem. Panel B: Embolisms caused by drought could be flushed and hydraulic function partially restored. Occlusion caused by tyloses could not be flushed which allow differentiation between drought induced and oak wilt induced dysfunction. Panel C: Example of a cross section of an oak wilt infected tree. Red staining depicts functional xylem. In red oaks, functional xylem is commonly located in the current and some of the last year's ring. Areas not stained in read are clogged by tyloses. Panel D: Oak wilt induced dysfunction also showed a spatial pattern typical of a pathogen that spreads from a point of infection to contiguous sections of the xylem whereas drought showed no spatial pattern.



DOI of published paper: https://doi.org/10.1073/pnas.2316164121

Appendix S10. Performance and output metrics for the Bayesian regression models. Variables described correspond to relative water content (RWC), loss of rehydration capacity (LRC), volumetric water content (VWC), electrolyte leakage (EL), maximum efficiency of photosystem II (FvFm), and water potential (WP). Metrics described correspond to the Gelman and Rubin's convergence diagnostic (Rhat) that stipulates convergence when values approach 1, the estimated threshold value (i.e., the days since the start of visible symptoms at which physiology starts to change in response to stress), and the 95% lower and upper confidence interval values for the estimated threshold. Values in bold indicate pre-visual detection (threshold confidence interval < 0).

Variable	Treatment	Rhat	Threshold value	Lower CI	Upper CI
Spectra-predicted Fv/Fm					
	OW	1.03	**-11.23**	**-22.80**	**-0.42**
	D	1.02	**-10.34**	**-22.60**	**-1.17**
	DxOW	1.02	**-7.33**	**-12.40**	**-2.66**
Spectra-predicted water potential					
	OW	1.01	-4.06	-8.43	0.72
	D	1.04	**-6.80**	**-11.99**	**-1.03**
	DxOW	1.03	**-4.95**	**-9.25**	**-1.46**
Spectra-predicted RWC					
	OW	1.02	-4.63	-10.42	1.20
	D	1.10	**-6.86**	**-12.31**	**-0.49**
	DxOW	1.04	**-4.60**	**-8.34**	**-0.81**
Spectra-predicted VWC					
	OW	1.01	-4.93	-11.63	0.91
	D	1.01	**-7.44**	**-12.68**	**-1.84**
	DxOW	1.07	**-4.47**	**-7.83**	**-0.76**
Spectra-predicted LRC					
	OW	1.01	**-6.26**	**-10.93**	**-0.90**
	D	1.01	**-7.54**	**-14.47**	**-1.87**
	DxOW	1.03	**-5.70**	**-9.85**	**-2.03**
Spectra-predicted EL					
	OW	1.01	-3.89	-10.71	2.74
	D	1.03	**-7.63**	**-13.89**	**-2.02**
	DxOW	1.02	**-5.35**	**-9.51**	**-1.34**