Supplementary dataset S4: Composite model summary (including individual models).

REMINDER: TWG = 1/3 x π x height x (diameter/2)^2 x wood density x glucose release

COMBINED MODEL FOR TWG	
R2	0.80
Q2	0.80

	Selected models to build the composite model for predicting TWG			
	Glucose release after pre-treatment (g.g-1) (Model type: GAM)	Wood density (g.mm-3) (Model type: Linear Model)	Stem height (mm) (Model type: Linear Model)	Stem diameter (mm) (Model type: Linear Model)
R2	0.79	0.81	0.99	0.85
Q2	0.49	0.77	0.51	0.69
Variables / importance*	ARAtms (smooth) — 3.632 RHAtms (linear) — -0.753 GALtms (linear) — 0.059 SG (linear) — 0.134 FibTanDiam (smooth) — 2.527 FibArea (linear) — 0.006 MOE Gpa (smooth) — 1.529 Interecept (if any): — -2.559	GLCtms — -0.00000193 FibRadDiam — -0.0000287 Wall thickness — 0.000331 Interecept (if any): — 0.00055	ARAtms — -860.25 FUCtms — -8070.78 XYLtms — -100.78 4-O-MeGlcAtms — -182.37 GLCtms — -84.69 G — 452.38 S — 359.28 H — 237.92 P — 540.62 CL — 6481.88 FibFrac — 12310.91 FibPop — -4.63 FibArea — -35.39 Interecept (if any): — 0	ARAtms — -4.6 RHAtms — -3.7 XYLtms — -0.51 MANtms — -0.74 4-O-MeGlcAtms — -1.18 GLCtms — -0.44 SG — -1.27 CL — 31.72 L — 1.73 Wall thickness — -1.63 Interecept (if any): — -41.11
Other information	Family: gamma Link function: log gamma = 1		The intercept was set to 0 by the user because the algorithm tended to overblow the importance of the intercept by returning	

* Coefficient for Linear Model and for linear terms in GAMs, number of degrees of freedom for smooth terms in GAM.

Orange color indicates that this trait affects both saccharification and biomass production.