Supplementary File 1

Retrospective recovery thresholds and nonlinear kinetic modelling of n-hexane Soxhlet extraction from *Moringa oleifera* and *Irvingia gabonensis*

Numerical methods, secondary cycle-based results and residual diagnostics

S1. Reproducible numerical analysis

The input is the 22 rounded endpoint mean pairs in main Table 1. Each point is the mean of two separate terminal extractions of fresh 15-g portions from one prepared batch per material. Individual duplicate times, masses and yields were not available. No zero-time observation, imputation, time correction, endpoint deletion or fabricated uncertainty estimate is included. All reported fits use the same input means.

Equivalent amplitude/scale parameterizations were fitted: $A \cdot x/(\tau+x)$ for Peleg, $A[1-\exp(-x/\tau)]$ for first-order, and $A\{1-\exp[-(x/\tau)^{\beta}]\}$ for Weibull. Here $A = Y\infty$; $K_1 = \tau/A$ and $K_2 = 1/A$ for Peleg, $k = 1/\tau$ for first-order, and $\alpha = \tau$ for Weibull. x is mean elapsed time or terminal cycle count. Fits minimize unweighted squared yield residuals.

SciPy least_squares used the trust-region reflective algorithm, a three-point finite-difference Jacobian, Jacobian-based parameter scaling and linear loss. Every fitted parameter had lower bound 10^{-10} and no finite upper bound; a 100% bound on A was not imposed. Initial values were $A = \max(Y) \times \{0.9, 1.1, 1.4\}$ and $\tau = \text{median}(x) \times \{0.15, 0.5, 1.0\}$; Weibull also used $\beta = \{0.6, 1.0, 1.4\}$. This gives 9, 9 and 27 starts per material and coordinate (180 attempts in total). ftol, xtol and gtol were 10^{-12}, with at most 10,000 function evaluations per start. Only converged finite fits were eligible; the smallest RSS was selected. Unsuccessful/non-finite attempts are logged and complete failure raises an error.

With $e_i = Y_i - \hat{Y}_i$, $RSS = \Sigma e_i^2$, $RMSE = \sqrt{(RSS/n)}$, $MAE = \Sigma|e_i|/n$, and $R^2 = 1-RSS/\Sigma(Y_i-\bar{Y})^2$. The working likelihood treats endpoint-mean errors as independent Gaussian errors of common unknown variance and treats mean coordinates as fixed. σ^2 is estimated as RSS/n. The full AICc is $n[\ln(2\pi)+1+\ln(RSS/n)] + 2K + 2K(K+1)/(n-K-1)$, with $K = p+1$, including residual variance. $p = 2$ or 3; $n = 10$ or 12 means. ΔAICc and weights are calculated within each material and coordinate. The conventional correction provides a conditional comparison of nonlinear fits, not evidence that its error assumptions hold.

Parameter intervals are not supplied: a conditional regression interval is mathematically possible, but mean-time uncertainty and the mean-response error model could not be assessed. The absent duplicate-level data prevent replicate-based dispersion estimates and separation of pure error from model lack of fit. Residual plots are diagnostics, not replicate error bars. The code records software versions, all initial values, optimization outcomes and fitted residuals in Supplementary File 2.

S2. Secondary analysis using terminal siphon-cycle count

Cycle count replaces mean elapsed time without changing endpoint means. This is an alternative descriptive coordinate; it is not an independent experiment or validation dataset.

Table S1. Complete cycle-based fit statistics.

Material	Model	RSS	RMSE	MAE	R^2	AICc	ΔAICc	Weight
M. oleifera	Peleg	12.042	1.097	0.890	0.9838	40.237	11.511	0.003
M. oleifera	First-order	3.809	0.617	0.493	0.9949	28.727	0.000	0.941
M. oleifera	Weibull	3.671	0.606	0.493	0.9951	34.358	5.631	0.056
I. gabonensis	Peleg	99.373	2.878	2.259	0.9698	68.422	7.552	0.020
I. gabonensis	First-order	52.962	2.101	1.601	0.9839	60.871	0.000	0.888
I. gabonensis	Weibull	52.211	2.086	1.549	0.9841	65.414	4.543	0.092

RSS is in squared percentage points; RMSE and MAE are in percentage points of yield. Full AICc includes residual variance (K = 3 for Peleg/first-order and 4 for Weibull). n = 10 or 12 endpoint means. AICc values are compared only within the same material and coordinate.

Table S2. Cycle-based parameter estimates and coordinates relative to the fitted asymptote.

Material	Model	Parameter estimates	Y_∞ (%)	c_{90} (cycles)	c_{95} (cycles)
M. oleifera	Peleg	$K_1 = 0.036053$; $K_2 = 0.018211$	54.91	17.82	37.62
M. oleifera	First-order	$k = 0.462895$	44.90	4.97	6.47
M. oleifera	Weibull	$\alpha = 2.1764$; $\beta = 0.9730$	45.13	5.13	6.72
I. gabonensis	Peleg	$K_1 = 0.024175$; $K_2 = 0.010137$	98.64	21.46	45.31
I. gabonensis	First-order	$k = 0.390331$	80.37	5.90	7.67
I. gabonensis	Weibull	$\alpha = 2.5410$; $\beta = 1.0315$	79.94	5.70	7.36

K_1: cycles per percentage point; K_2: inverse percentage point; k: cycle^{-1}; α: cycles; β: dimensionless. c_{90} and c_{95} are continuous model coordinates at 90% and 95% of Y_∞, not the discrete empirical thresholds in main Table 2. The fitted asymptote is not verified total extractable oil.

S3. Residual diagnostics

Figures S1–S4 assess the time-based fits. The positive *I. gabonensis* residual at 34 min is retained for all three models. Peleg's residual pattern in this material cautions against relying on a high R^2 or an extrapolated plateau alone. These short sets of endpoint means do not establish homoscedasticity or Gaussian errors. No measured within-pair variance is inferred from the plots.

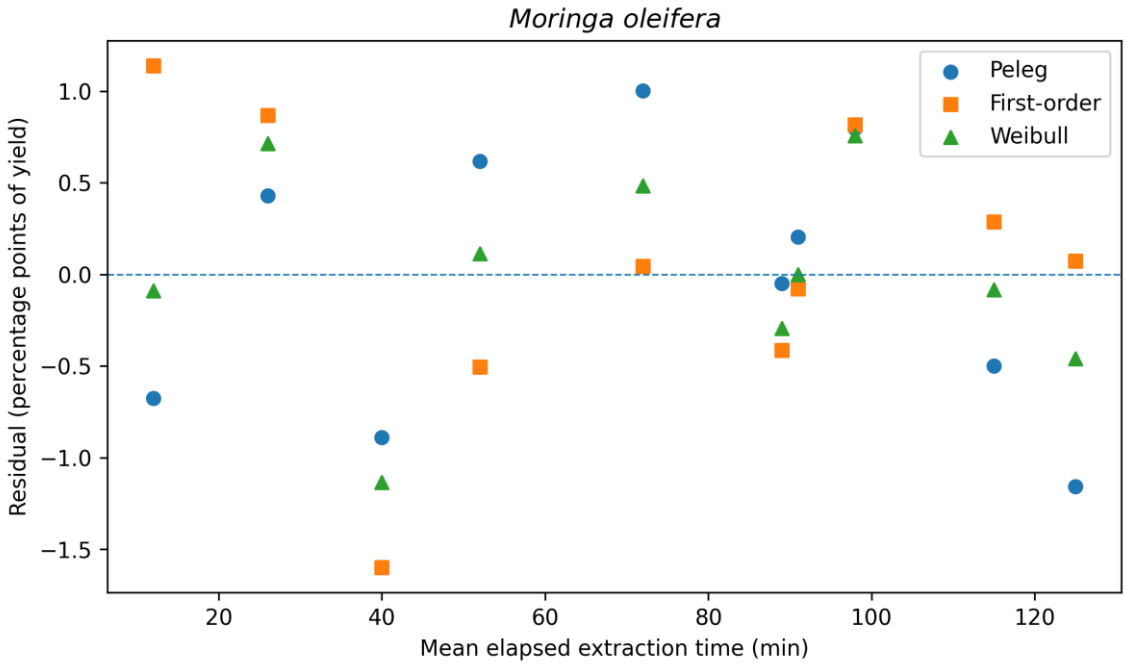


Figure S1. Residuals for *Moringa oleifera* versus mean elapsed time. Residuals are observed endpoint mean minus fitted yield, in percentage points. Symbols distinguish models, not individual replicates; the dashed line marks zero. All endpoint means are retained.

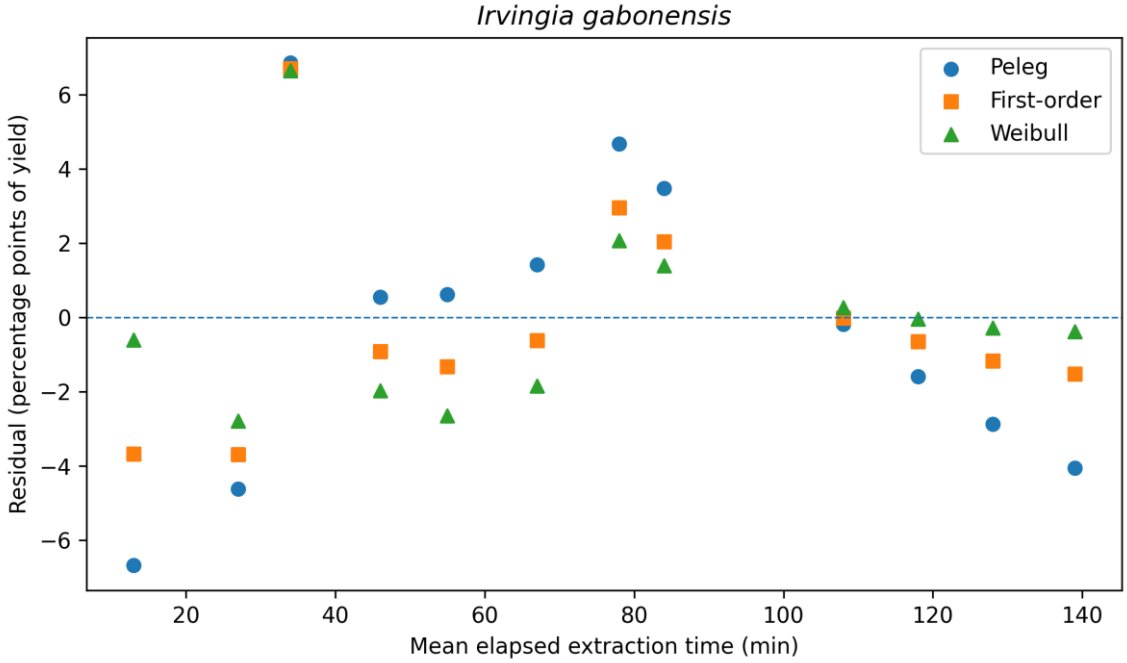


Figure S2. Residuals for *Irvingia gabonensis* versus mean elapsed time. Residuals are observed endpoint mean minus fitted yield, in percentage points. Symbols distinguish models, not individual replicates; the dashed line marks zero. All endpoint means are retained.

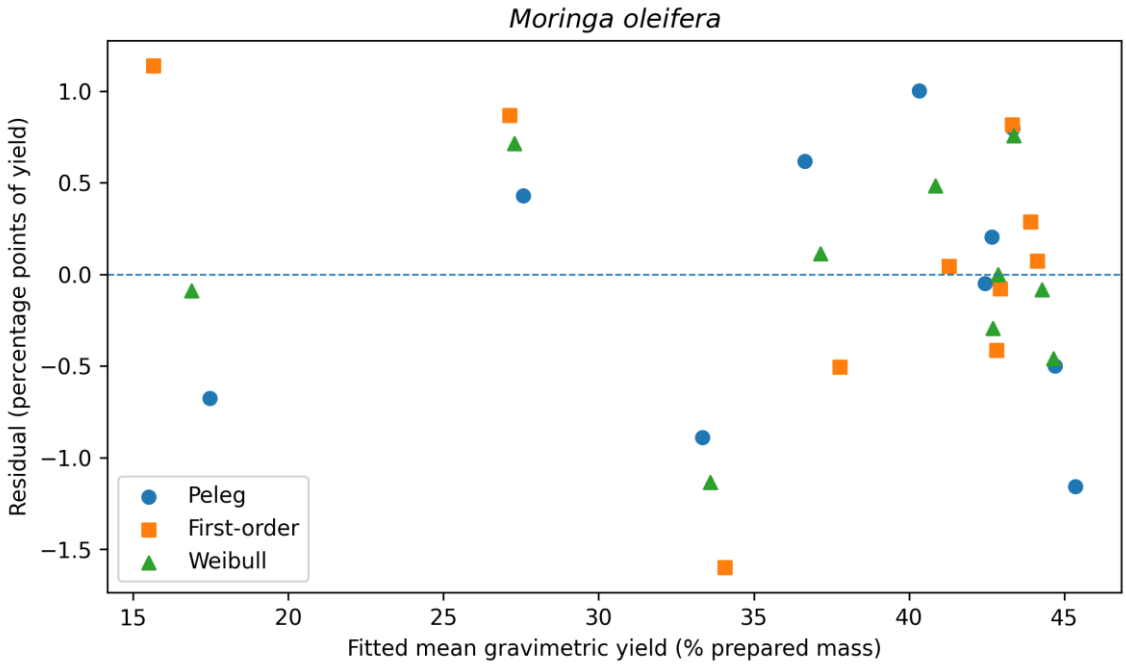


Figure S3. Residuals for *Moringa oleifera* versus fitted mean yield. Residuals are observed endpoint mean minus fitted yield, in percentage points. Symbols distinguish models, not individual replicates; the dashed line marks zero. All endpoint means are retained.

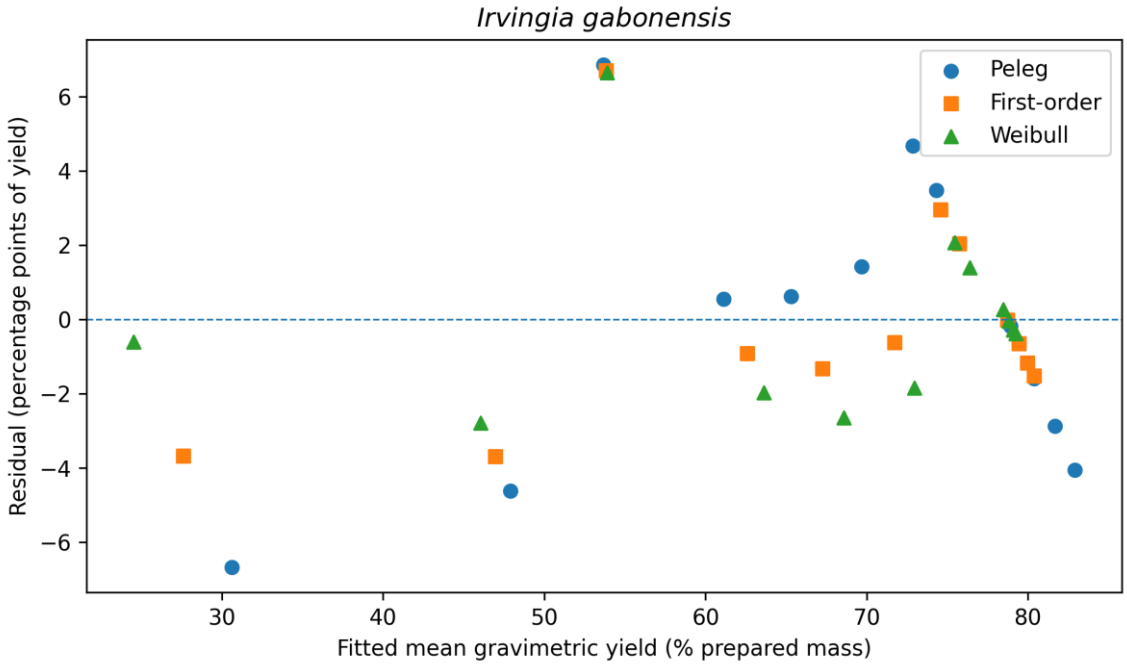


Figure S4. Residuals for *Irvingia gabonensis* versus fitted mean yield. Residuals are observed endpoint mean minus fitted yield, in percentage points. Symbols distinguish models, not individual replicates; the dashed line marks zero. All endpoint means are retained.