Event-Conditioned Robust Gradient Matching for Lotka–Volterra Parameter Estimation: A Variance-Based Framework for Informative Threshold Selection

Ved Lakshminarayan

I would like to thank my supervisor, Professor Wenying Shou, for their mentorship and Dr Roman Ulises Zapien Campos for their invaluable guidance and support in shaping the project.

Abstract

Parameter estimation in nonlinear ordinary differential equations is commonly performed by repeatedly solving the governing equations and minimising discrepancy between simulated and observed trajectories. This work develops an alternative event-conditioned estimator for the classical Lotka–Volterra predator–prey system. Rather than fitting the entire trajectory during primary inference, the proposed method selects informative population threshold-crossing events and estimates the model parameters from local state and derivative information at those events.

Noisy population observations are smoothed and differentiated using Savitzky–Golay local polynomial regression. Candidate events are defined by upward and downward crossings of population quantiles. Because the Lotka–Volterra vector field is linear in its unknown parameters, the local prey and predator equations can then be expressed as two linear-in-parameter regression systems. We evaluate candidate threshold sets using normalised design-matrix determinants and condition numbers.

A central analytical result is that the raw determinant of the prey design matrix admits the exact representation

$$\det(X_x^\top X_x) = S_x^2 \operatorname{Var}_w(y), \qquad S_x = \sum_i x_i^2, \qquad w_i = \frac{x_i^2}{S_x},$$

while the predator design satisfies

$$\det(X_y^\top X_y) = S_y^2 \operatorname{Var}_v(x), \qquad S_y = \sum_i y_i^2, \qquad v_i = \frac{y_i^2}{S_y}.$$

Thus prey parameter identifiability requires variation in predator abundance at selected events, whereas predator parameter identifiability requires variation in prey abundance. After column normalisation, the determinants become monotone functions of weighted coefficients of variation.

Parameters are estimated using positivity-constrained Huber iteratively reweighted least squares, with the Rousseeuw–Croux Q_n estimator used to robustly scale residuals. Uncertainty is quantified by resampling complete oscillatory cycles.

For a synthetic Lotka–Volterra experiment with true parameter vector $(1, 0.1, 0.075, 1.5)$, the method selected three quantile levels, 0.10, 0.15, and 0.65, retaining 50 events. The estimated parameters were

$$(0.98431,\ 0.098667,\ 0.074950,\ 1.494014),$$

corresponding to a relative parameter RMSE of 0.01049. All 20 validation experiments completed successfully, with a maximum relative parameter RMSE of 0.03248. The results

demonstrate that state-dependent event selection can retain strong parameter information while substantially compressing the trajectory used for primary estimation.

1 Introduction

Ordinary differential equation models are widely used to represent biological interaction mechanisms, but estimating their parameters from noisy measurements can be difficult. A standard approach is *trajectory matching*: a candidate parameter vector is proposed, the differential equation is numerically integrated, and the resulting trajectory is compared with observed data. Optimisation then repeatedly modifies the parameters and resolves the ODE.

Although such procedures use all available observations, they can be computationally expensive and sensitive to initial conditions, phase alignment, observation noise and numerical optimisation. Furthermore, not every observation necessarily contributes equally to distinguishing the unknown parameters.

An alternative family of methods exploits local derivatives. Varah [3] proposed a two-stage procedure in which observed data were first approximated using smoothing splines, after which derivatives of the smooth representation were matched to the differential equation. Modern approaches often refer to this strategy as *gradient matching*. Related methods range from sparse regression approaches such as SINDy [6] to weak-form parameter inference such as WENDy [7].

The present work considers an additional question:

> *Can a small number of dynamically informative events contain enough local information to estimate the parameters of an oscillatory system?*

For the Lotka–Volterra predator–prey system, threshold crossings provide natural state-dependent events. Instead of evaluating the gradient-matching equations at every time point, candidate events are generated whenever either population crosses one of several quantile levels. A subset of these quantiles is then selected according to the geometry of the resulting parameter-design matrices.

The central theoretical contribution is an explicit connection between the determinant of those design matrices and weighted cross-species variance. This gives a direct biological interpretation to identifiability. Separating prey growth from predation requires the predator population to vary across the selected prey observations; separating predator gain from mortality analogously requires variation in prey abundance.

The resulting framework combines event-based data reduction, design-matrix geometry, robust gradient matching and cycle-level uncertainty quantification.

2 Mathematical Model

Let $x(t) > 0$ denote prey abundance and $y(t) > 0$ predator abundance. The classical Lotka–Volterra equations are

$$\dot{x} = \alpha x - \beta xy, \tag{1}$$
$$\dot{y} = \delta xy - \gamma y. \tag{2}$$

The unknown parameter vector is

$$\boldsymbol{\theta} = (\alpha, \beta, \delta, \gamma)^\top.$$

The parameter α represents prey growth in the absence of predators, β represents prey loss due to predator encounters, δ describes predator growth resulting from encounters, and γ represents predator mortality in the absence of prey.

Although Equations (1)–(2) are nonlinear in the states through the term xy, they are linear in the unknown parameters. This distinction is fundamental to the proposed estimator.

2.1 Why level information alone cannot identify the time scale

Write the system as

$$\dot{\boldsymbol{z}} = F(\boldsymbol{z}; \boldsymbol{\theta}), \qquad \boldsymbol{z} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

If all four parameters are multiplied by a common constant $c > 0$,

$$F(\boldsymbol{z}; c\boldsymbol{\theta}) = cF(\boldsymbol{z}; \boldsymbol{\theta}).$$

Suppose $\boldsymbol{z}_{\boldsymbol{\theta}}(t)$ is the solution associated with $\boldsymbol{\theta}$ and define

$$\boldsymbol{w}(t) = \boldsymbol{z}_{\boldsymbol{\theta}}(ct).$$

Then

$$\frac{d\boldsymbol{w}}{dt} = cF(\boldsymbol{w}; \boldsymbol{\theta}) = F(\boldsymbol{w}; c\boldsymbol{\theta}).$$

Therefore,

$$\boxed{\boldsymbol{z}_{c\boldsymbol{\theta}}(t) = \boldsymbol{z}_{\boldsymbol{\theta}}(ct).}$$

The two parameter vectors $\boldsymbol{\theta}$ and $c\boldsymbol{\theta}$ therefore generate the same orbit in (x, y) state space but traverse it at different speeds.

Consequently, data containing only

$$\{\text{crossed level, species, direction, order}\}$$

cannot identify the common temporal scale of the parameters. Derivative or timing information is necessary. The present method uses local derivatives estimated independently from the observed trajectory.

3 State and Derivative Estimation

Suppose observations are available at times t_i,

$$\mathcal{D} = \left\{ (t_i, x_i^{\mathrm{obs}}, y_i^{\mathrm{obs}}) \right\}_{i=1}^{n}.$$

When the time spacing is not sufficiently uniform, the observations are first represented on an equally spaced time grid using piecewise-linear interpolation.

Savitzky–Golay smoothing [1] is then performed separately for each species. Around time t_i, a local polynomial

$$P_i(\tau) = a_{0,i} + a_{1,i}\tau + \cdots + a_{p,i}\tau^p, \qquad \tau = t - t_i,$$

is fitted by least squares to observations inside a moving window.

At the centre of the local window,

$$\widehat{x}(t_i) = a_{0,i},$$

and

$$\widehat{\dot{x}}(t_i) = a_{1,i}.$$

An analogous calculation provides $\widehat{y}(t_i)$ and $\widehat{\dot{y}}(t_i)$.

The implementation uses a cubic polynomial, $p = 3$, and searches the candidate window lengths

$$21,\ 31,\ 41,\ 51,\ 61.$$

A window h is scored using a combination of data mismatch and local roughness,

$$J(h) = \frac{1}{n}\sum_i \left(\frac{z_i - \widehat{z}_{h,i}}{s_z} \right)^2 + 0.015 \frac{1}{n-2} \sum_i \left(\frac{\Delta^2 \widehat{z}_{h,i}}{s_h} \right)^2.$$

The first term discourages excessive smoothing, whereas the second discourages an unnecessarily rough estimated trajectory. The window minimising $J(h)$ is selected independently for prey and predator.

4 Event Construction

Candidate quantiles are

$$\mathcal{Q} = \{0.10, 0.15, 0.20, \ldots, 0.90\}.$$

For species x, quantile q corresponds to the population threshold

$$L_{x,q} = Q_q(\widehat{x}),$$

and analogously

$$L_{y,q} = Q_q(\widehat{y}).$$

For a smoothed signal u_i and a level L, define

$$d_i = u_i - L.$$

An upward crossing is detected when

$$d_i < 0, \qquad d_{i+1} \geq 0,$$

and a downward crossing when

$$d_i > 0, \qquad d_{i+1} \leq 0.$$

The exact crossing usually lies between sampled observations. Assuming local linearity,

$$f = \frac{L - u_i}{u_{i+1} - u_i}$$

gives its fractional position between t_i and t_{i+1}. Thus

$$\tau = (1 - f)t_i + f t_{i+1}.$$

The same fraction is used to interpolate all required quantities,

$$x(\tau), \qquad y(\tau), \qquad \dot{x}(\tau), \qquad \dot{y}(\tau).$$

Each event therefore contains

$$e_j = (\tau_j, x_j, y_j, \dot{x}_j, \dot{y}_j, q_j, s_j, d_j)\,,$$

where s_j denotes the crossed species and d_j the crossing direction.

Events lying in the first or final 4% of the time-series indices are discarded because derivative estimation is least reliable near the boundaries.

5 Local Parameter Regression

At event i, the prey equation gives

$$\dot{x}_i = \alpha x_i - \beta x_i y_i.$$

Collecting N selected events,

$$\boldsymbol{b}_x = X_x \boldsymbol{\theta}_x + \boldsymbol{\varepsilon}_x,$$

where

$$\boldsymbol{b}_x = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_N \end{bmatrix}, \qquad X_x = \begin{bmatrix} x_1 & -x_1 y_1 \\ x_2 & -x_2 y_2 \\ \vdots & \vdots \\ x_N & -x_N y_N \end{bmatrix},$$

and

$$\boldsymbol{\theta}_x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Likewise,

$$\boldsymbol{b}_y = X_y \boldsymbol{\theta}_y + \boldsymbol{\varepsilon}_y,$$

with

$$X_y = \begin{bmatrix} x_1 y_1 & -y_1 \\ x_2 y_2 & -y_2 \\ \vdots & \vdots \\ x_N y_N & -y_N \end{bmatrix}, \qquad \boldsymbol{\theta}_y = \begin{bmatrix} \delta \\ \gamma \end{bmatrix}.$$

Importantly, the responses $\dot{x}_i$ and $\dot{y}_i$ are the Savitzky–Golay derivative estimates. They are not generated from the Lotka–Volterra model. The regression therefore asks which parameter values make the known Lotka–Volterra vector-field form agree most closely with the independently estimated local slopes.

6 Variance Interpretation of Identifiability

The determinant of the Gram matrix $X^\top X$ quantifies whether the two design columns span substantially different directions.

6.1 Prey parameters

For the prey design,

$$X_x^\top X_x = \begin{bmatrix} \sum_i x_i^2 & -\sum_i x_i^2 y_i \\ -\sum_i x_i^2 y_i & \sum_i x_i^2 y_i^2 \end{bmatrix}.$$

Therefore,

$$\det(X_x^\top X_x) = \left(\sum_i x_i^2\right)\left(\sum_i x_i^2 y_i^2\right) - \left(\sum_i x_i^2 y_i\right)^2. \tag{3}$$

Define

$$S_x = \sum_i x_i^2$$

and the normalised weights

$$w_i = \frac{x_i^2}{S_x}.$$

Because

$$\sum_i w_i = 1,$$

the weighted predator mean is

$$\mu_{y,w} = \sum_i w_i y_i,$$

and the weighted variance is

$$\mathrm{Var}_w(y) = \sum_i w_i (y_i - \mu_{y,w})^2.$$

Using

$$\mathrm{Var}_w(y) = \sum_i w_i y_i^2 - \left(\sum_i w_i y_i\right)^2,$$

Equation (3) becomes

$$\boxed{\det(X_x^\top X_x) = S_x^2 \,\mathrm{Var}_w(y).}$$

Thus, for α and β to be distinguishable, predator abundance must vary across the selected events.

If all selected events satisfy

$$y_i = c,$$

then

$$\mathrm{Var}_w(y) = 0,$$

so

$$\det(X_x^\top X_x) = 0.$$

The equation reduces to

$$\dot{x}_i = x_i(\alpha - c\beta),$$

meaning that only the combined quantity

$$\alpha - c\beta$$

can be identified.

6.2 Predator parameters

Similarly,

$$X_y^\top X_y = \begin{bmatrix} \sum_i x_i^2 y_i^2 & -\sum_i x_i y_i^2 \\ -\sum_i x_i y_i^2 & \sum_i y_i^2 \end{bmatrix}.$$

Define

$$S_y = \sum_i y_i^2, \qquad v_i = \frac{y_i^2}{S_y}.$$

Then

$$\boxed{\det(X_y^\top X_y) = S_y^2 \,\mathrm{Var}_v(x).}$$

Consequently, distinguishing predator gain δ from predator mortality γ requires variation in prey abundance at the selected events.

6.3 Normalised determinant

The implementation normalises each design column before calculating the selection criterion. If

$$\widetilde{X} = \begin{bmatrix} \boldsymbol{c}_1/\|\boldsymbol{c}_1\| & \boldsymbol{c}_2/\|\boldsymbol{c}_2\| \end{bmatrix},$$

then

$$G = \widetilde{X}^\top \widetilde{X} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

where ρ is the cosine of the angle between the columns.

Therefore,

$$\boxed{\det(G) = 1 - \rho^2.}$$

For the prey design, the variance identity yields

$$\det(G_x) = \frac{\mathrm{Var}_w(y)}{\mathbb{E}_w[y^2]}.$$

Since

$$\mathbb{E}_w[y^2] = \mu_{y,w}^2 + \mathrm{Var}_w(y),$$

defining

$$CV_w(y) = \frac{\sqrt{\mathrm{Var}_w(y)}}{\mu_{y,w}}$$

gives

$$\boxed{\det(G_x) = \frac{CV_w(y)^2}{1 + CV_w(y)^2}.}$$

Likewise,

$$\boxed{\det(G_y) = \frac{CV_v(x)^2}{1 + CV_v(x)^2}.}$$

Thus the normalised determinant is a monotone measure of weighted *relative cross-species variation*. This provides a direct statistical and biological interpretation of the selection criterion.

The eigenvalues of G are

$$\lambda_\pm = 1 \pm |\rho|,$$

giving condition number

$$\boxed{\kappa(G) = \frac{1 + |\rho|}{1 - |\rho|}.}$$

As $|\rho| \to 1$, the two parameter effects become nearly indistinguishable and the condition number diverges.

7 Threshold-Level Selection

For a candidate quantile set $\mathcal{L}$, the prey and predator normalised Gram matrices are constructed from all events generated by those quantiles.

The joint score is

$$\boxed{I(\mathcal{L}) = \log\det(G_x(\mathcal{L})) + \log\det(G_y(\mathcal{L})).}$$

This criterion resembles a D-optimal determinant criterion, although because the design columns are normalised it should more precisely be interpreted as a *normalised parameter-separation score* rather than a full Fisher information matrix.

Quantiles are added greedily. Starting from the empty set, the candidate quantile that produces the highest joint score is added at each stage. The search continues up to eight levels.

A candidate design is admissible if it contains at least two levels, at least 12 events, and satisfies

$$\kappa_x \le 10^5, \qquad \kappa_y \le 10^5.$$

The chosen K is the smallest admissible number of levels for which the relative improvement from adding another level is below 0.05.

8 Robust Parameter Estimation

Direct derivative estimates may contain isolated errors arising from measurement noise, smoothing, interpolation, or local model mismatch. Ordinary least squares gives such observations quadratic influence. The final estimator therefore uses Huber robust regression [2].

First, the columns of each design matrix are scaled to unit Euclidean norm to improve numerical conditioning. A positivity-constrained least squares estimate supplies the initial parameter vector.

At iteration m,

$$\boldsymbol{r}^{(m)} = \boldsymbol{b} - X\boldsymbol{\theta}^{(m)}$$

denotes the residual vector.

Residual scale is estimated using the Rousseeuw–Croux Q_n estimator [5]. For residuals $r_1, \dots, r_N$, form all pairwise distances

$$d_{ij} = |r_i - r_j|, \qquad i < j.$$

Let

$$h = \left\lfloor \frac{N}{2} \right\rfloor + 1, \qquad k = \binom{h}{2}.$$

The implementation uses

$$\boxed{Q_n = 2.2219\, d_{(k)},}$$

where $d_{(k)}$ is the kth smallest pairwise distance.

The standardised residual is

$$u_i = \frac{r_i}{Q_n}.$$

Huber's loss is

$$\rho_c(u) = \begin{cases} \dfrac{1}{2}u^2, & |u| \le c, \\ c|u| - \dfrac{1}{2}c^2, & |u| > c, \end{cases}$$

with

$$c = 1.345.$$

The corresponding IRLS weights are

$$w_i^H = \begin{cases} 1, & |u_i| \le c, \\ \dfrac{c}{|u_i|}, & |u_i| > c. \end{cases}$$

The next parameter estimate solves

$$\boxed{\boldsymbol{\theta}^{(m+1)} = \arg\min_{\boldsymbol{\theta} \ge \boldsymbol{\theta}_{\min}} \sum_i w_i^H \left(b_i - X_i\boldsymbol{\theta}\right)^2.}$$

All four parameters are constrained to be positive, with numerical lower bound 10^{-10}.

The iteration terminates when

$$\|\boldsymbol{\theta}^{(m+1)} - \boldsymbol{\theta}^{(m)}\|_2 \le 10^{-9}\left(1 + \|\boldsymbol{\theta}^{(m)}\|_2\right).$$

9 Cycle-Block Bootstrap

Threshold events belonging to the same predator–prey oscillation are not statistically independent. A single noisy oscillation can influence several crossings. Resampling event rows independently would therefore destroy within-cycle dependence.

The trajectory is partitioned into cycles using peaks of the smoothed prey population. Let

$$\mathcal{C}_1, \ldots, \mathcal{C}_M$$

denote the resulting event blocks.

For each bootstrap replicate, M cycle labels are drawn with replacement and all events belonging to each selected cycle are retained. The complete parameter estimator is then rerun,

$$\widehat{\boldsymbol{\theta}}^{*(b)} = T\left(\mathcal{C}_{I_1^{(b)}}, \ldots, \mathcal{C}_{I_M^{(b)}}\right).$$

This is conceptually related to block-bootstrap methods for dependent observations [4].

The bootstrap estimates the conditional variability of the parameter fit given the extracted events and chosen levels. It does not currently repeat smoothing, crossing detection, or level selection inside each replicate.

10 Synthetic Experiment

A controlled numerical experiment was used because the true parameters were then known exactly.

The generating parameters were

$$\boldsymbol{\theta}_{\text{true}} = (1,\ 0.10,\ 0.075,\ 1.50)^{\top}$$

with initial state

$$(x_0, y_0) = (10, 5).$$

The clean model was simulated over

$$0 \le t \le 24$$

at 800 equally spaced times.

Independent Gaussian measurement noise was added to each species. The standard deviation for each species was set to 0.012 times its clean trajectory range.

For the principal experiment, 400 cycle-bootstrap replicates were used.

11 Results

11.1 Smoothing and event selection

Both prey and predator selected a Savitzky–Golay window length of 21.

The greedy selection sequence began with quantile 0.15. At $K = 2$, the 0.65 quantile was added, and at $K = 3$, quantile 0.10 was added.

The selected set was therefore

$$\boxed{\mathcal{L} = \{0.10,\ 0.15,\ 0.65\}.}$$

These thresholds generated 50 retained events distributed across four detected oscillation cycles.

Table 1 shows the relevant stages of the selection procedure.

Table 1: Information-based level selection.

K	Events	$I(K)$	$\log\det G_x$	$\log\det G_y$	Gain
1	16	-3.2497	-1.1600	-2.0898	–
2	32	-2.7640	-1.1574	-1.6067	0.1495
3	50	-2.4330	-1.0457	-1.3872	0.1198
4	68	-2.3158	-0.9787	-1.3371	0.0481

Because the improvement from three to four levels was

$$0.0481 < 0.05,$$

the stopping rule selected

$$\boxed{K = 3}.$$

The selected design condition numbers were

$$\boxed{\kappa_x = 9.274, \qquad \kappa_y = 13.943,}$$

well below the allowed threshold of 10^5.

Exponentiating the log determinants gives

$$\det(G_x) = 0.3514, \qquad \det(G_y) = 0.2498.$$

The corresponding weighted coefficients of variation were

$$CV_w(y) = 0.736, \qquad CV_v(x) = 0.577.$$

Thus the selected events contained substantial cross-species relative variation for both regressions.

11.2 Parameter estimates

The robust event estimator returned

$$\boxed{\widehat{\boldsymbol{\theta}} = \begin{bmatrix} 0.984310 \\ 0.098667 \\ 0.074950 \\ 1.494014 \end{bmatrix}.}$$

The true values and relative errors are shown in Table 2.

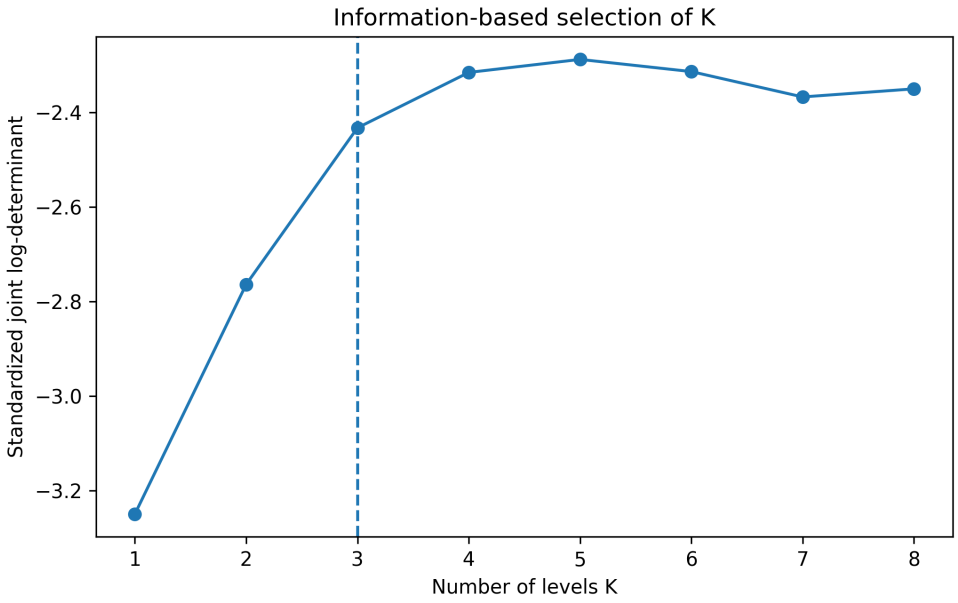

Figure 1: Joint normalised design score as the number of retained threshold levels increases. The dashed line marks the selected value $K = 3$.

Table 2: True and estimated Lotka–Volterra parameters.

Parameter	True	Estimate	Relative error (%)
α	1.000000	0.984310	-1.569
β	0.100000	0.098667	-1.333
δ	0.075000	0.074950	-0.067
γ	1.500000	1.494014	-0.399

The combined relative parameter RMSE,

$$\mathrm{RMSE}_{\mathrm{rel}} = \sqrt{\frac{1}{4}\sum_{j=1}^{4}\left(\frac{\widehat{\theta}_j - \theta_j}{\theta_j}\right)^2},$$

was

$$\boxed{\mathrm{RMSE}_{\mathrm{rel}} = 0.01049.}$$

Thus, the typical relative parameter discrepancy was approximately 1.05%.

Forward integration using the estimated parameter vector produced a trajectory normalised RMSE of

$$\boxed{0.05664.}$$

This forward calculation was not used to obtain the primary estimate; it served as an external trajectory-level validation of the local event fit.

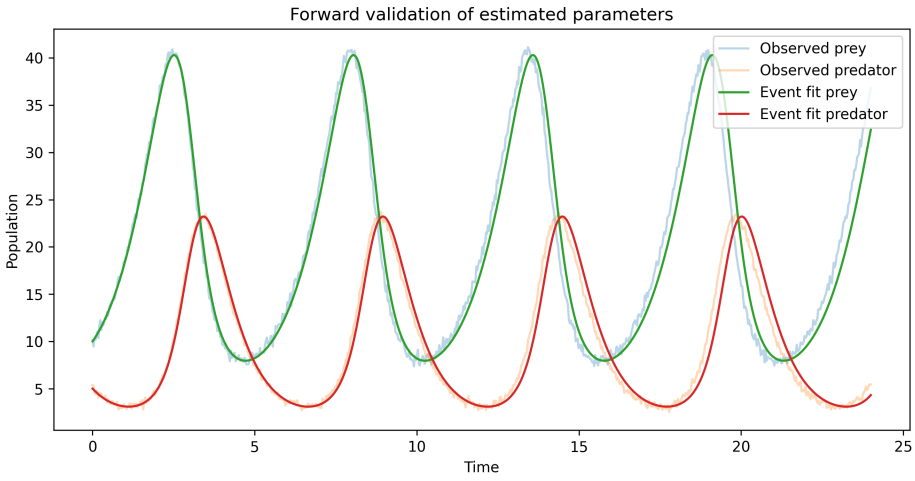

Figure 2: Forward simulation using the event-estimated parameter vector compared with the noisy observed trajectory.

11.3 Robust weighting

The fitted Q_n residual scales were approximately

$$Q_{n,x} = 1.412, \qquad Q_{n,y} = 0.995.$$

Huber weighting reduced the influence of 9 of the 50 prey derivative residuals and 6 of the 50 predator residuals. The smallest resulting weights were approximately 0.534 and 0.581, respectively.

This confirms that the robust estimator was not equivalent to ordinary least squares in the principal experiment: several locally anomalous events were meaningfully downweighted rather than removed.

11.4 Bootstrap uncertainty

The cycle-block bootstrap results are reported in Table 3.

Table 3: Cycle-block bootstrap parameter uncertainty.

Parameter	Bootstrap mean	Bootstrap SD	2.5%	97.5%
α	0.98251	0.01677	0.94163	1.01091
β	0.098475	0.000894	0.096511	0.100046
δ	0.074875	0.000996	0.072643	0.076895
γ	1.49240	0.02066	1.45045	1.53158

All four true generating parameter values were contained within their corresponding 95% percentile intervals.

The bootstrap should nevertheless be interpreted cautiously because only four independent oscillatory blocks were available. Increasing the number of bootstrap resamples does not increase the number of independent biological cycles.

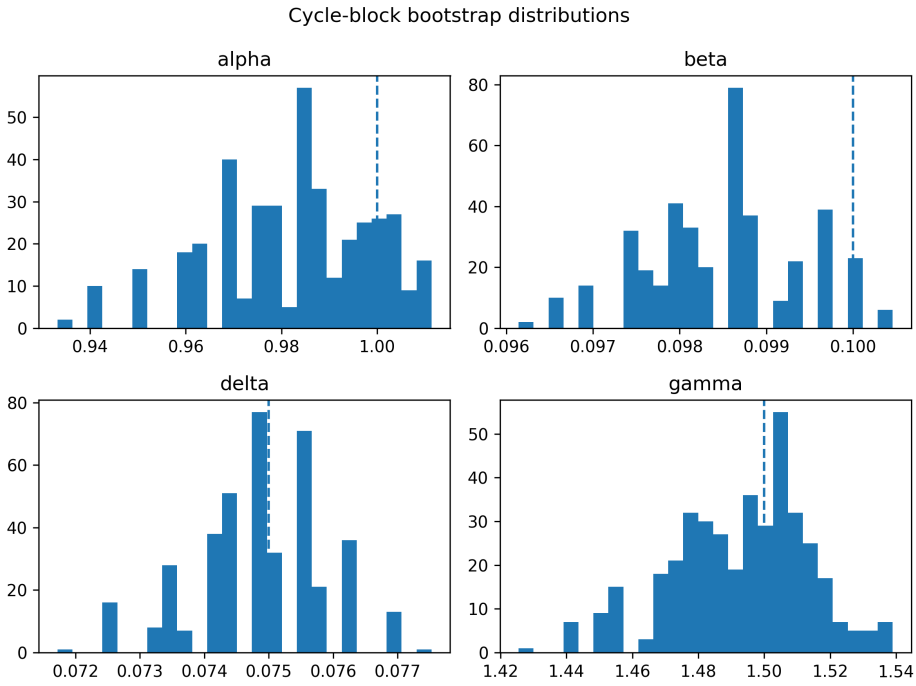

Figure 3: Cycle-block bootstrap distributions. Dashed vertical lines denote the true generating parameters.

11.5 Validation experiments

The estimator was additionally evaluated through 20 complete validation runs spanning five categories:

1. five random noise realisations;
2. five observation-noise amplitudes;
3. three candidate smoothing-window families;
4. four initial conditions;
5. three repeated identical runs.

All 20 runs completed successfully.

Across the complete validation set, the median relative parameter RMSE was

$$\boxed{0.01049},$$

while the largest observed value was

$$\boxed{0.03248}.$$

Table 4 summarises the grouped performance.

Noise sensitivity followed the expected pattern. With no added observation noise, relative parameter RMSE was approximately

Table 4: Summary of validation experiments.

Validation group	Runs	Median relative RMSE	Maximum relative RMSE
Random noise seed	5	0.01049	0.02067
Noise level	5	0.01049	0.03248
Smoothing windows	3	0.01049	0.01355
Initial condition	4	0.01383	0.02258
Reproducibility	3	0.01049	0.01049

$$1.85 \times 10^{-4},$$

whereas at noise fractions 0.005, 0.012, 0.020, and 0.035 the relative RMSE values were approximately

$$0.00223, \quad 0.01049, \quad 0.01976, \quad 0.03248,$$

respectively.

The deterioration with noise is consistent with the principal weakness of direct gradient matching: numerical derivative estimation becomes less reliable as observational noise increases.

12 Discussion

The results support the central hypothesis that a comparatively small set of state-dependent events can retain sufficient information for accurate parameter estimation in the classical Lotka–Volterra system.

The most important theoretical result is not simply that a determinant can be maximised, but that the determinant has a direct variance interpretation. For the prey equation,

$$\dot{x} = x(\alpha - \beta y),$$

the parameter effects become indistinguishable when y is approximately constant. The identity

$$\det(X_x^\top X_x) = S_x^2 \operatorname{Var}_w(y)$$

makes this exact. The determinant is zero precisely when weighted predator variation vanishes. The predator equation has the symmetric property

$$\det(X_y^\top X_y) = S_y^2 \operatorname{Var}_v(x).$$

This interpretation explains why event location matters. A large number of measurements concentrated in one narrow state-space region may contain less parameter-separation information than a smaller number of measurements spanning dynamically different states.

Column normalisation further separates this issue from raw population scale. The resulting determinant,

$$\det(G_x) = \frac{CV_w(y)^2}{1 + CV_w(y)^2},$$

depends only on weighted relative variation. Consequently, the level selection mechanism is effectively searching for threshold sets that produce high relative variation in the opposing population while maintaining numerical conditioning.

The main experiment selected only three of the seventeen candidate quantiles. The first level, 0.15, alone produced a relatively poorly conditioned predator design, with predator condition number approximately 30.3. Adding 0.65 reduced this to 17.9, and adding 0.10 reduced it further to 13.9. The fourth level produced only a 4.8% relative gain in the joint score and was therefore rejected by the predefined stopping rule. This illustrates the intended principle of the method: additional events are retained only while they provide appreciable new parameter separation.

The resulting parameter accuracy was high for the tested synthetic system. All four estimates were within 1.6% of the true values, and the combined relative RMSE was approximately 1.05%. The robust weights were active for several events, indicating that the result was not merely a conventional least-squares fit disguised as a robust procedure.

The approach belongs conceptually to the broader family of gradient matching methods. Varah [3] similarly proposed smoothing observed trajectories, differentiating the smooth representation and estimating differential-equation parameters from those derivatives. The principal distinction here is the event-conditioned design: regression is intentionally restricted to threshold crossings chosen for parameter separability.

SINDy [6] also performs derivative regression on nonlinear functions of the state, but its objective is structural discovery from a candidate function library. In contrast, the present work assumes the Lotka–Volterra structure and estimates only its coefficients.

WENDy [7] provides an especially relevant modern comparison. It avoids direct numerical differentiation by transforming the governing equations into a weak form and explicitly addresses an errors-in-variables structure. The present approach instead retains direct local derivatives but introduces state-dependent event selection. A future comparison between event-conditioned gradient matching and event-conditioned weak-form estimation would therefore be particularly informative.

13 Limitations and Future Work

Several limitations require emphasis.

First, derivative estimation is the principal noise-sensitive stage. Savitzky–Golay smoothing reduces noise amplification but cannot remove it entirely. The validation study showed progressively increasing parameter error with increasing observation noise.

Second, the design variables x_i and y_i and the derivative responses $\dot{x}_i$ and $\dot{y}_i$ are all derived from the same smoothed observations. The regression is therefore not a classical fixed-design model with error-free predictors. An errors-in-variables treatment would provide a more complete statistical description.

Third, exact threshold-crossing times are not observed from discrete samples. Linear interpolation assumes approximately linear motion over each sampling interval. Very sparse sampling would weaken this assumption.

Fourth, the Q_n implementation uses the asymptotic Gaussian consistency multiplier 2.2219 but does not currently apply an explicit finite-sample correction. This is a minor issue for large event counts but could matter for very small selected sets.

Fifth, the cycle bootstrap conditions on the selected levels and extracted events. It therefore quantifies parameter uncertainty given the selected design rather than complete pipeline uncertainty. A full-pipeline bootstrap would resample cycles before smoothing and rerun smoothing, event detection, level selection and regression in every replicate.

Finally, only four oscillatory cycles were available in the principal synthetic experiment. The resulting bootstrap intervals are useful as a stability diagnostic but should not be interpreted as highly accurate large-sample confidence intervals.

Natural extensions include weak-form or integral event estimation, explicit propagation of smoothing uncertainty, level-selection stability analysis, stochastic predator–prey models, and testing on experimental ecological data.

14 Conclusion

This study developed an event-conditioned robust gradient-matching framework for estimating the four parameters of the classical Lotka–Volterra predator–prey system.

The method replaces repeated full-trajectory ODE fitting during primary estimation with a sequence of local operations: smoothing and differentiation, threshold-crossing extraction, information-based level selection and robust constrained regression.

The key mathematical result is

$$\boxed{\det(X_x^\top X_x) = S_x^2 \operatorname{Var}_w(y), \qquad \det(X_y^\top X_y) = S_y^2 \operatorname{Var}_v(x).}$$

Thus parameter identifiability is governed directly by cross-species variation at the chosen event locations. After design normalisation, the criterion becomes a monotone function of weighted relative variation.

In the corrected synthetic experiment, only three threshold quantiles were required. Fifty selected events yielded estimates

$$(0.98431,\ 0.098667,\ 0.074950,\ 1.494014)$$

for true parameters

$$(1,\ 0.1,\ 0.075,\ 1.5),$$

with a relative parameter RMSE of approximately 1.05%. All 20 validation experiments completed successfully, and the maximum relative parameter RMSE across those experiments remained below 3.25%.

The results suggest that informative state-dependent events can provide a compact and interpretable representation of local dynamical information. More broadly, the variance derivation demonstrates that the usefulness of a dynamical observation is determined not simply by how many observations are collected, but by whether those observations separate the effects of the unknown parameters.

References

[1] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964. doi:10.1021/ac60214a047.

[2] P. J. Huber, "Robust Estimation of a Location Parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964. doi:10.1214/aoms/1177703732.

[3] J. M. Varah, "A Spline Least Squares Method for Numerical Parameter Estimation in Differential Equations," *SIAM Journal on Scientific and Statistical Computing*, vol. 3, no. 1, pp. 28–46, 1982. doi:10.1137/0903003.

[4] H. R. Künsch, "The Jackknife and the Bootstrap for General Stationary Observations," *The Annals of Statistics*, vol. 17, no. 3, pp. 1217–1241, 1989. doi:10.1214/aos/1176347265.

[5] P. J. Rousseeuw and C. Croux, "Alternatives to the Median Absolute Deviation," *Journal of the American Statistical Association*, vol. 88, no. 424, pp. 1273–1283, 1993. doi:10.1080/01621459.1993.10476408.

[6] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering Governing Equations from Data by Sparse Identification of Nonlinear Dynamical Systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016. doi:10.1073/pnas.1517384113.

[7] D. M. Bortz, D. A. Messenger, and V. Dukic, "Direct Estimation of Parameters in ODE Models Using WENDy: Weak-Form Estimation of Nonlinear Dynamics," *Bulletin of Mathematical Biology*, vol. 85, article 110, 2023. doi:10.1007/s11538-023-01208-6.