Supplementary Material II: Experiments - Tables and Figures

Contents

Introduction

The supplementary material contains additional figures and tables to the paper "Parametric Estimation of Ordinary Differential Equations with Orthogonality Conditions". It gives more information about the results of the Monte Carlo simulations for the comparison of the TS, NLS and OC estimators. Indeed, for the evaluation of estimator accuracy, we compute the Absolute Relative Error (ARE) defined by

$$ARE = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} \frac{\left|\theta^* - \widehat{\theta}_i\right|}{|\theta^*|}.$$

Additionally, we compute the weighted Orthogonal Condition estimators obtained by minimizing the criterion $Q_{n,L}^{W}(\theta) = \boldsymbol{e}_L\left(\hat{\phi}, \theta\right)^{\top} W \boldsymbol{e}_L\left(\hat{\phi}, \theta\right)$, where the weight matrix computed by the IRWOC algorithm, as introduced and motivated in section 5 of *Supplementary Material I.* We also compute the coverage probabilities of the estimator (NLS, OC and OC with an optimally weighted weight matrix): we look at the influence of the sample size n and of the noise level confidence ellipses derived from the asymptotic (Gaussian) approximation.

The supplementary material contains additional tables and figures for the models introduced in the paper:α-pinene, Riccati, and also for the real data example: Influenza virus growth and Blowfly populations dynamics. Moreover, we have tested our approach with the "FitzHugh-Nagumo model" for neuron dynamics. This model is chosen as it is an example of 2D nonlinear ODE with nonlinear dependence in the parameters (with a periodic solution). This model was introduced by Ramsay et al. as a benchmark for ODE estimation. We recall then briefly the model, and we compare the 3 approaches discussed in the paper. Concerning the Riccati equation, we plot the objective functions minimized by NLS and OC, in order to explain the difference of the NLS and OC estimators when the change-point time is not known.

Comments

From the simulations, we can check that the confidence sets given by NLS are too stringent, and NLS overestimates the coverage probabilities, as the (Monte-Carlo) estimated coverage probabilities are quite smaller than 95% (for all simulated models: α-pinene, Riccati, FitzHugh-Nagumo). On the contrary for OC, we can check the reliability of the derived confidence sets, as the estimated coverage probability is around 95% (or even higher). This is partly due to the estimated variance of OC that is higher than the NLS (as we can see by comparing the sum of estimated parameters variances $Tr\left(V\left(\hat{\theta}\right)\right)$). Hence, this also confirms the quality of the Gaussian approximation derived from the asymptotic analysis of the paper.

The Optimally Weighted OC was introduced in order to ameliorate the variance of the OC estimator, in order to reduce the size of the Confidence sets while preserving the coverage probabilities. Although this version should improve on the unweighted OC estimator, this is not always true in practice, and the "Optimal" estimator can be worse. This is mainly due to the fact that we approximate the variance by linearization technics, and we do not control the error propagation of this approximation in the IRWOC algorithm. Hence, we are not sure that we obtain the optimal matrix. Moreover the control of the nonsingularity of the variance matrix can force us to reduce the number of orthogonal conditions, which can induce a loss of inference power. Nevertheless, the optimal weighting approaches improve upon the simple OC estimator for α-pinene and FitzHugh-Nagumo; in addition the optimal weighting estimator gives the best estimate for the Influenza Virus Growth model.

1 Linear ODE: α-*pinene*

1.1 Known initial condition

$\times 10^{-2}$	*MSE*							*Tr* $\left(V(\hat{\theta})\right)$				
(n,σ)	TS	OC	OCopt	OC,0	OC,0 opt	NLS		OC	OCopt	OC,0	OC,0 opt	NLS
$(400,3)$	0.72	0.05	0.04	0.04	0.03	0.02		0.04	0.04	0.04	0.05	0.02
$(400,8)$	2.28	0.22	0.27	0.25	0.22	0.10		0.95	1.91	1.20	2.23	0.12
$(200,3)$	1.19	0.27	0.33	0.30	0.32	0.03		0.09	0.20	0.13	0.24	0.03
$(200,8)$	2.95	0.44	0.51	0.37	0.47	0.18		2.66	4.96	2.68	5.48	0.27
$(50,3)$	2.39	0.27	0.27	0.26	0.26	0.16		1.37	2.75	1.58	3.31	0.16
$(50,8)$	4.54	1.03	0.94	0.93	0.89	0.68		7.96	8.20	7.27	9.13	1.68

$\times 10^{-2}$	*ARE*					
(n,σ)	TS	OC	OCopt	OC,0	OC,0 opt	NLS
$(400,3)$	105.85	20.45	19.83	18.99	17.64	11.61
$(400,8)$	213.69	47.90	49.48	50.77	45.29	28.65
$(200,3)$	150.53	37.16	40.36	36.38	37.40	16.58
$(200,8)$	235.62	57.50	60.64	58.54	59.97	38.93
$(50,3)$	220.59	51.19	52.41	51.13	51.72	34.49
$(50,8)$	283.02	112.52	110.44	107.21	107.77	76.52

Table 1: MSE, Asymptotic Variance & ARE for α-pinene model with known Initial Condition

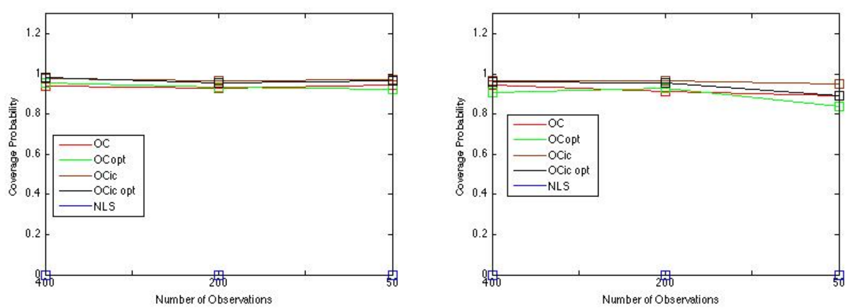

Figure 1: Coverage Probabilities for the 95% Confidence Ellipse for α-pinene model with known Initial Condition

1.2 Unknown initial condition

$\times 10^{-2}$	*MSE*							$Tr\left(V(\hat{\theta})\right)$				
(n,σ)	TS	OC	OCopt	OC,1	OC,1 opt	NLS		OC	OCopt	OC,1	OC,1 opt	NLS
$(400,3)$	0.25	0.11	0.25	0.11	0.52	0.07		0.10	0.11	0.10	0.09	0.06
$(400,8)$	1.07	0.85	0.85	0.56	0.73	0.50		1.06	0.94	0.82	0.45	0.61
$(200,3)$	0.6	0.37	0.50	0.23	0.34	0.14		0.25	0.25	0.20	0.17	0.14
$(200,8)$	1.64	1.42	1.18	0.83	1.68	1.34		2.36	2.30	1.64	0.50	1.54
$(50,3)$	1.33	1.31	1.18	0.80	1.04	0.69		1.63	1.46	1.02	0.39	0.76
$(50,8)$	3.64	2.11	1.91	1.79	3.44	1.96		5.34	4.35	2.20	0.81	4.38

$\times 10^{-2}$	*ARE*					
(n,σ)	TS	OC	OCopt	OC,1	OC,1 opt	NLS
$(400,3)$	60.13	32.66	39.51	32.58	56.48	24.13
$(400,8)$	127.82	86.64	88.53	79.02	112.08	63.52
$(200,3)$	88.75	52.84	59.02	47.48	75.10	34.70
$(200,8)$	158.17	117.44	113.69	108.70	177.77	98.85
$(50,3)$	138.31	99.87	97.36	89.22	121.73	69.27
$(50,8)$	247.57	161.97	164.31	165.55	214.20	144.83

Table 2: MSE, Asymptotic Variance & ARE for α-pinene model with unknown initial conditions

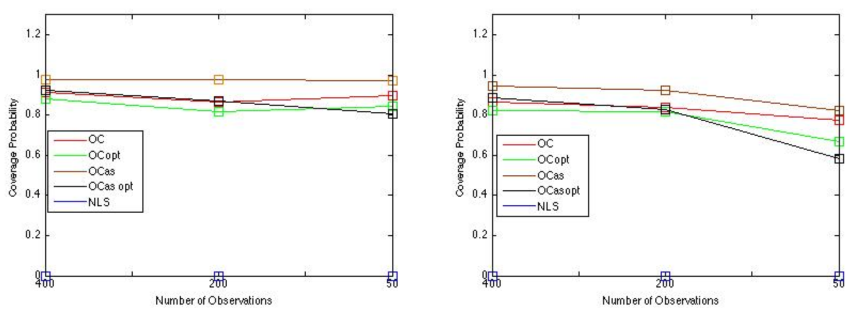

Figure 2: Coverage Probabilities for the 95% Confidence Ellipse for α-Pinene with unknown Initial Conditions

2 Nonlinear ODEs

2.1 Ricatti equation

We recall that the asymptotic criterion for the nonlinear least squares is

$$\begin{aligned} R^*_{NLS}(\theta, x_0) &= E_{T,Y}\left[\|Y - \phi(T;\theta,x_0)\|^2\right] \\ &= \sigma^2 + \int_0^1 \|\phi(t;\theta,x_0) - \phi(t;\theta^*,x_0^*)\|^2 \pi(t)dt \end{aligned}$$

whereas the asymptotic criterion for OC is

$$Q^*_L(\theta) = \sum_{\ell=1}^{L} e_\ell(\phi^*, \theta)^2.$$

When the Ricatti equation is simply $\dot{x} = ax^2 + c\sqrt{t}$, with unknown a and c, the criterion R_{NLS} is nonlinear in θ (no closed-form), whereas the $Q^*_L(\theta)$ is a simple quadratic form, which is simple to optimize. In that case, we can check (by Monte-Carlo simulations) that the optimization of the empirical counterpart of $R_{NLS}(\theta, x_0)$ gives better results (when we start close to θ^*) than $Q^*_L(\theta)$. Nevertheless, when the parametrized ODE is $\dot{x} = ax^2 + c\sqrt{t} - d'c\mathbb{1}_{[T_r;14]}$, the statistical inverse problem becomes much harder to solve. This can be seen by plotting the criterion function $R^*_{NLS}(\theta, x_0)$ which is a rough function, that makes it difficult to optimize. At the contrary, the function $Q^*_L(\theta)$ is quite smooth and exhibits (locally) only one minimum at θ^*. In order to visualize the 2 objective functions, we fix $a = a^*$ and $c = c^*$ and we consider $d' \in [0, 3]$ and $T_r \in [10, 14]$ (for NLS, we have also to fix $x_0 = x_0^* = -1$). We recall that we have $d'^* = 2$ and $T_r^* = 11$.

2.1.1 Known change-point time T_r

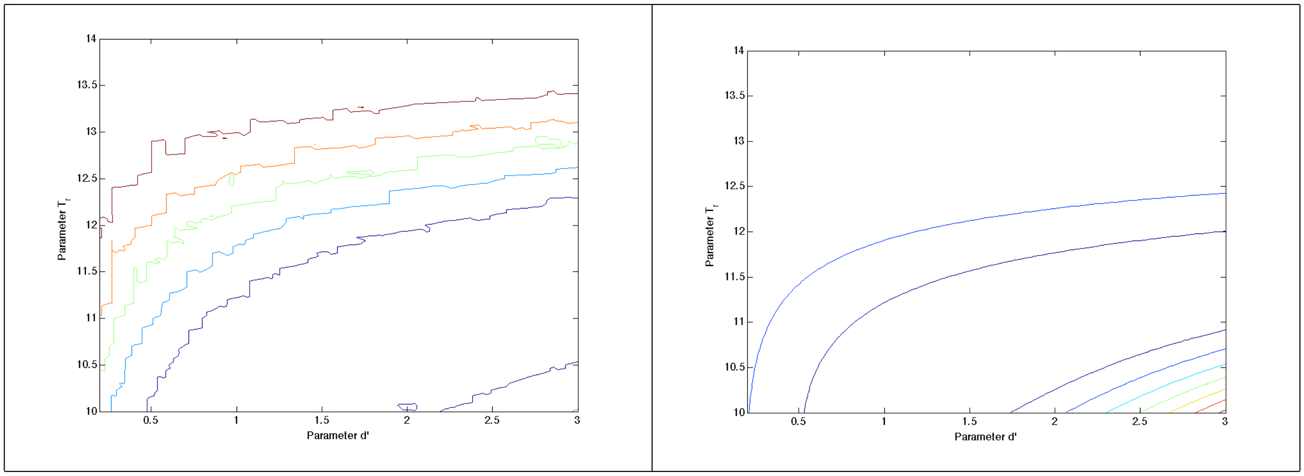

Figure 3: Contour plots for Asymptotic Statistical Criterions in Ricatti ODE: (Left) NLS criterion $R^*_{NLS}(d', T_r)$; (Right) OC criterion $Q^*_L(d', T_r)$
$(d'^* = 2, T_r^* = 11)$.

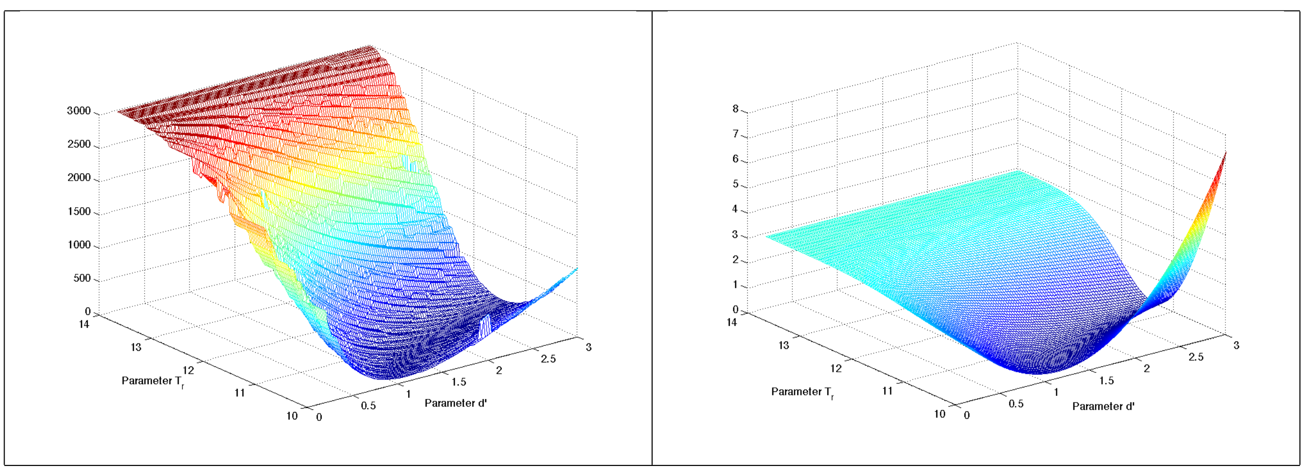

Figure 4: Contour plots for Asymptotic Statistical Criterions in Ricatti ODE: (Left) NLS criterion $R^*_{NLS}(d', T_r)$; (Right) OC criterion $Q^*_L(d', T_r)$, $(d'^* = 2, T_r^* = 11)$

$\times 10^{-2}$	MSE				ARE				$Tr\left(V\left(\hat{\theta}\right)\right)$		
(n,σ)	TS	OC	OCopt	NLS	TS	OC	OCopt	NLS	OC	OCopt	NLS
$(400, 0.2)$	0.18	0.27	0.51	0.58	21.75	16.47	26.10	18.54	1.76	1.44	0.10
$(400, 0.4)$	0.78	1.21	1.32	0.94	45.03	31.98	42.82	36.25	2.56	1.80	0.38
$(200, 0.2)$	0.33	0.87	1.04	0.57	27.09	26.95	39.63	25.81	2.85	2.07	0.25
$(200, 0.4)$	1.12	2.69	2.62	1.12	53.99	43.17	54.04	45.10	5.64	2.98	0.98
$(50, 0.2)$	1.03	1.30	2.08	1.54	45.53	41.50	49.98	49.12	4.70	5.39	1.00
$(50, 0.4)$	3.80	4.43	5.06	3.94	77.08	73.69	73.30	70.96	8.89	8.76	4.08

Table 3: MSE , ARE & $Tr\left(V\left(\hat{\theta}\right)\right)$ for Parameter estimation for Ricatti Equation with known T_r

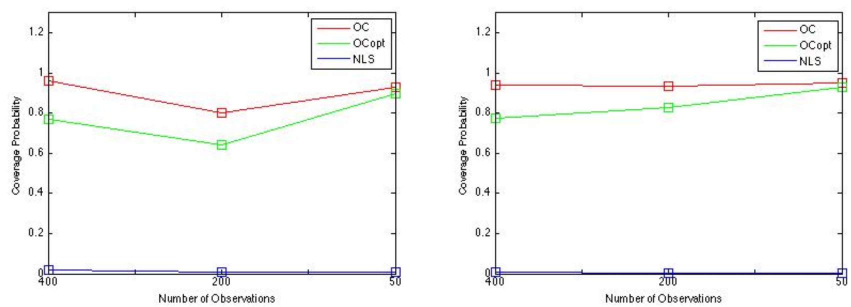

Figure 5: Coverage Probabilities for the 95% Confidence Ellipse for Ricatti model with known T_r

2.1.2 Unknown change-point time T_r

$\times 10^{-2}$	$MSE(\widehat{a})$		$MSE(\widehat{c})$		$MSE(\widehat{d'})$		$MSE(\widehat{T_r})$	
(n,σ)	OC	OCopt	OC	OCopt	OC	OCopt	OC	OCopt
$(400, 0.2)$	0.09	0.07	0.00	0.00	2.54	2.35	1.39	1.15
$(400, 0.4)$	0.29	0.26	0.01	0.01	4.27	3.63	3.54	2.95
$(200, 0.2)$	0.21	0.21	0.00	0.00	4.08	4.44	3.18	3.21
$(200, 0.4)$	0.61	0.55	0.01	0.01	11.96	12.80	6.93	7.88
$(50, 0.4)$	0.64	0.48	0.02	0.02	11.20	14.06	14.25	15.92
$(50, 0.4)$	0.77	0.89	0.01	0.02	17.18	18.63	19.40	18.43

$\times 10^{-2}$	MSE		ARE		$Tr\left(V(\hat{\theta})\right)$	
(n,σ)	OC	OCopt	OC	OCopt	OC	OCopt
$(400, 0.2)$	4.01	3.57	30.90	27.67	3.97	3.53
$(400, 0.4)$	8.11	6.85	56.12	54.42	8.02	6.62
$(200, 0.2)$	7.47	7.87	44.56	45.11	7.35	7.77
$(200, 0.4)$	19.51	21.23	78.91	73.47	18.94	20.82
$(50, 0.2)$	26.10	30.48	84.33	81.99	5.14	8.61
$(50, 0.4)$	37.36	37.96	98.67	106.39	9.49	12.90

Table 4: MSE ARE & Sum Empirical Variance for Parameter estimation for Ricatti with unknown T_r

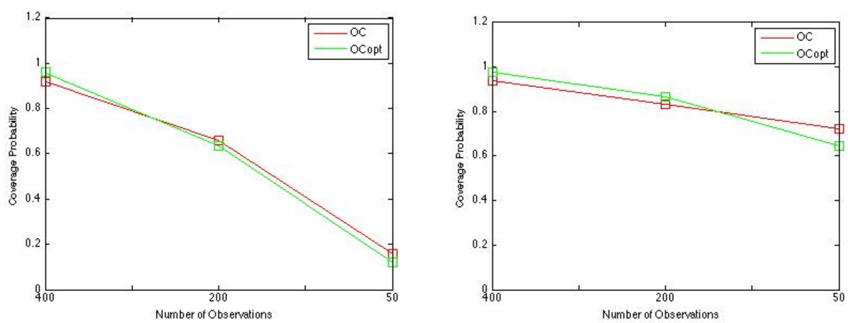

Figure 6: Coverage Probabilities for the 95% Confidence Ellipse for Ricatti model with unknown T_r

2.2 FitzHugh-Nagumo equation

The FitzHugh-Nagumo is a nonlinear two-dimensional ODE introduced for modeling neurons. For well-chosen sets of parameters and initial conditions, it exhibits a periodic behavior, with typical oscillations corresponding to a limit cycle.

$$\left\{ \begin{array}{lll} \dot{V} & = & c\left(V - \frac{V^3}{3} + R\right) \\ \dot{R} & = & -\frac{1}{c}\left(V - a + bR\right) \end{array} \right. \tag{1}$$

The true parameters are $a^* = b^* = 0.2$ and $c^* = 3$ and $x_0 = (V_0, R_0) = (-1, 1)$, and are taken from Ramsay et al (2007) where it was introduced as a benchmark for parameter estimation in ODEs. Due to the periodicity of the FitzHugh-Nagumo solution for this parameter set we use the sine basis for the test function and we choose the best number of orthogonal condition between 5 or 6.

$\times 10^{-2}$	$MSE(\widehat{a})$				$MSE(\widehat{b})$				$MSE(\widehat{c})$			
(n, σ)	TS	OC	OCopt	NLS	TS	OC	OCopt	NLS	TS	OC	OCopt	NLS
$(400, 0.15)$	0.01	0.01	0.01	0.00	0.09	0.08	0.07	0.05	3.54	0.94	0.12	0.01
$(400, 0.3)$	0.04	0.04	0.03	0.01	0.36	0.31	0.26	0.17	23.29	4.70	0.49	0.04
$(200, 0.15)$	0.02	0.02	0.01	0.01	0.20	0.18	0.14	0.08	9.80	2.01	0.25	0.02
$(200, 0.3)$	0.06	0.08	0.05	0.01	0.72	0.68	0.62	0.33	49.58	9.97	0.97	0.07
$(50, 0.15)$	0.06	0.07	0.11	0.02	0.77	0.73	1.03	0.29	168.01	36.54	3.07	0.06
$(50, 0.3)$	0.25	0.28	0.32	0.06	1.96	2.14	3.13	1.10	173.38	58.76	4.47	0.25

$\times 10^{-2}$	MSE				ARE				$Tr\left(V(\hat{\theta})\right)$		
(n, σ)	TS	OC	OCopt	NLS	TS	OC	OCopt	NLS	OC	OCopt	NLS
$(400, 0.15)$	3.64	1.04	0.20	0.07	21.60	18.12	15.02	11.20	1.09	0.17	0.02
$(400, 0.3)$	23.69	5.05	0.77	0.22	46.77	36.06	28.69	20.30	3.99	0.68	0.09
$(200, 0.15)$	10.02	2.21	0.40	0.11	32.73	26.12	20.98	14.64	1.95	0.3	0.05
$(200, 0.3)$	50.36	10.73	1.64	0.41	66.98	52.63	42.82	28.76	6.62	1.29	0.16
$(50, 0.15)$	168.84	37.35	4.20	0.37	80.47	58.13	56.39	26.75	4.16	0.77	0.18
$(50, 0.3)$	175.59	61.17	7.91	1.41	122.27	107.11	103.35	53.40	9.82	3.29	0.70

Table 5: MSE , ARE & Sum Asymptotic Variance for Parameter estimation for FitzHugh-Nagumo model

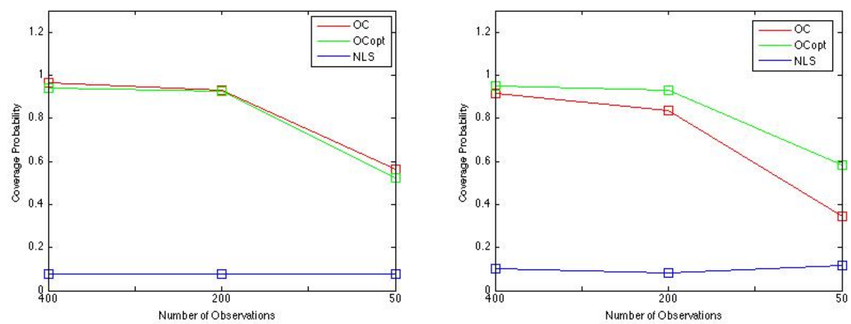

Figure 7: Ellipse set for parameter estimation for FitzHugh-Nagumo model

Results are presented in table 5 and Figure 7. We can see that NLS estimators is the best among all. Among the non-parametric estimators, the optimal one is OC opt, both in term of MSE and ARE. A parameter by parameter study shows us IRWOC algorithms dramatically improve the accuracy of the estimation of c.

3 Real data analysis

3.1 Influenza virus growth and migration model

	$\hat{\theta}_3^{OC}$	$\hat{\theta}_4^{OC}$	$\hat{\theta}^{NLS}$	$\tilde{\theta}^{ref}$
ρ_m	2.9e-5	2.7e-5	1.5e-5	$1.6e-5$
ρ_s	4.1e-5	4.7e-5	4.1e-5	$4.5e-5$
δ_l	2.0	3.4	3.7	3.96
γ_{ms}	0.39	0.35	0.15	0.157
γ_{sl}	0.72	0.81	0.47	0.49
RMSE	13.5	13.9	9.0	9.5

	$\hat{\theta}_3^{OC}$		$\hat{\theta}_4^{OC}$		$\hat{\theta}^{NLS}$	
	Low. Bound	Up. Bound	Low. Bound	Up. Bound	Low. Bound	Up. Bound
ρ_m	2.1e-5	3.7e-5	1.9e-5	3.4e-5	0.7e-0.5	2.4e-0.5
ρ_s	0.7e-5	7.4e-5	0.9e-5	8.4e-5	3.4e-0.5	4.8e-0.5
δ_l	-1.11	5.21	-0.28	7.21	2.59	4.93
γ_{ms}	0.27	0.50	0.24	0.46	0.03	0.26
γ_{sl}	-0.10	1.55	-0.14	1.76	0.39	0.55

Table 6: Estimates, RMSE and the 95% confidence intervals for different L and estimators.

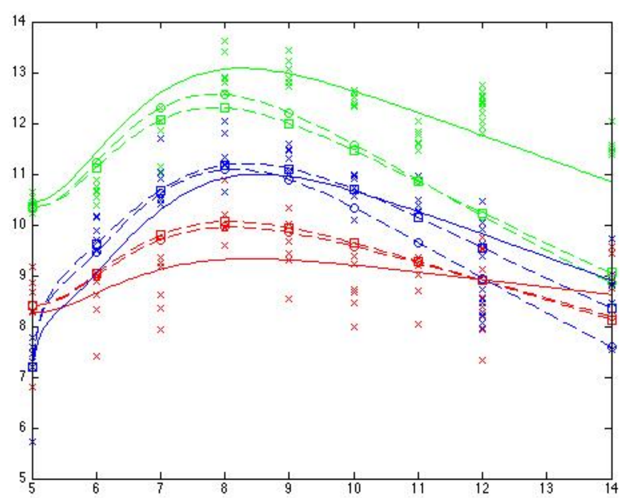

Figure 8: Influenza: Estimated curves for X_1 (red), X_2 (green), X_3 (blue); $\times$: observations, $\square$: solution for $\hat{\theta}_1^{OC}$, $\circ$: solution for $\hat{\theta}_2^{OC}$, solid line: solution with $\hat{\theta}^{NLS}$.

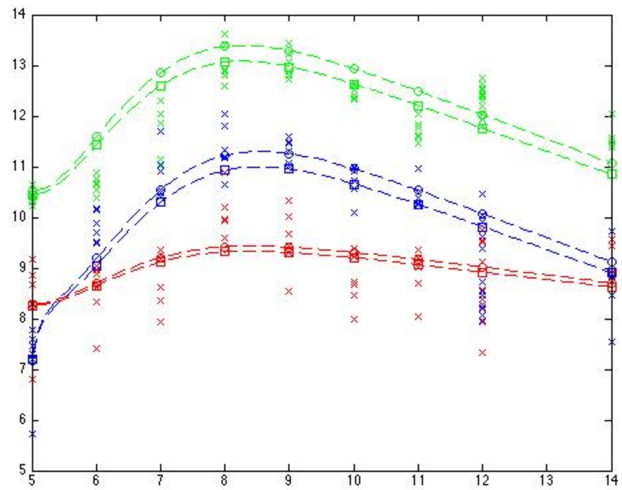

Figure 9: Influenza: Estimated curves for X_1 (red), X_2 (green), X_3 (blue); $\square$ solution obtained with OC+NLS, $\circ$ solution obtained with $\tilde{\theta}^{ref}$.

3.2 Blowfly models

	$L = 11$	$L = 9$	$L = 12$
P	7.81	7.52	7.91
N_0	381.8	385.9	377.7
δ	0.154	0.153	0.154
RSSE	1.7136e+03	1.7557e+03	1.7990e+03

	$L = 11$		$L = 9$		$L = 12$	
O.C	Low. Bound	Up. Bound	Low. Bound	Up. Bound	Low. Bound	Up. Bound
P	5.80	9.81	5.64	9.40	5.0416	10.77
N_0	303.62	459.94	306.59	465.38	289.36	465.98
δ	0.10	0.20	0.11	0.19	0.10	0.20

Table 7: Estimates, RSSE and 95% confidence intervals for different L