

### 5.4.1 Topological constraints on the GLV

Fitting the full GLV model to population dynamics never produces estimates of interaction strength that are exactly equal to zero. Therefore in some cases it is desirable to constrain the interaction topology which is used to fit the dynamics, in order to compare how the GLV fit performs without certain links. The method for conducting this constrained fit is outlined here. We reproduce the analytic solution for the GLV fit to species  $i$  (equation 5.36):

$$\hat{J}_i = Y G_i^T \left( G_i G_i^T \right)^{-1},$$

where the matrices are defined as

$$Y_i = \begin{pmatrix} \hat{y}_{i1} & \hat{y}_{i2} & \cdots & \hat{y}_{iM} \end{pmatrix} \in \mathbb{R}^{1 \times M},$$

$$\hat{J}_i = \begin{pmatrix} r_i & J_{i1} & J_{i2} & \cdots & J_{iN} \end{pmatrix} \in \mathbb{R}^{1 \times (N+1)},$$

$$G_i = \begin{pmatrix} f_{i1} & f_{i2} & \cdots & f_{iM} \\ g_{i11} & g_{i12} & \cdots & g_{i1M} \\ g_{i21} & g_{i22} & \cdots & g_{i2M} \\ \vdots & \vdots & \ddots & \vdots \\ g_{iN1} & g_{iN2} & \cdots & g_{iNM} \end{pmatrix} \in \mathbb{R}^{(N+1) \times M}.$$

We can constrain the species  $j$  with which species  $i$  is allowed to interact in the fitted model. By removing the element  $\hat{y}_{ij}$  from  $Y_i$ , and the  $j^{th}$  row from  $G_i$ , we effectively remove the  $j^{th}$  species from the error minimisation. In this case we obtain an estimate  $\hat{J}_i \in \mathbb{R}^{1 \times (N)}$  (rather than  $\mathbb{R}^{1 \times (N+1)}$ ). As many species as required can be removed from the error minimisation, and when the full estimated interaction matrix  $\hat{J}$  is reconstructed we insert zeros into the corresponding matrix entries.

### 5.4.2 Properties of the data generator

Before applying the inference method to *data streams* derived from IBM simulations, we consider certain properties of the IBM. This model has yet to be used to simulate communities with a small number of species. In previous chapters we have studied simulated communities of no less than 30 species. The simulation of smaller communities requires some parameter adjustment in order to obtain dynamics that are amenable to our goal of estimating the strength of species interactions. We also consider the functional forms generated by the model, in order to determine

the appropriateness of fitting a GLV model to the dynamics. Specifically we study the intrinsic growth functions of basal and non-basal species, and the functional response of non-basal predators.

Unlike the ODE models used as the *data generators* in section 5.3, the IBM does not allow us to calculate analytic forms for the interaction strengths (using the metric  $\alpha$ , defined in section 5.2.4). Rather the interaction strengths in the IBM emerge from the local interactions between individuals during the simulation. In this sense the use of the IBM as *data generator* represents a step towards *realism* in testing the inference method. However, the unavailability of analytic forms for the interaction strengths raises the question of how to evaluate the performance of the inference method. The simplest check, and the most fundamental concern, is to determine if the fitted GLV parameters correctly identify which species are prey, which species are predators, and which species interact. That is, does the method identify the correct interaction network topology. Further to this we are interested in determining how accurate the GLV estimates of interaction strength are, compared to some other measure of interaction strength derived from the IBM output. One possibility is to use the metric IS (equation (2.20)), which has been used hitherto in the thesis. However this metric only quantifies the per-capita effect on the prey population, per-capita of the predator. Therefore IS is analogous to the GLV estimate  $\hat{J}_{01}$ . In section ?? we develop a method, related to the way IS is calculated<sup>4</sup>, that allows us to quantify all inter-specific interaction strengths, and a combined measure of the intrinsic parameters for each species.

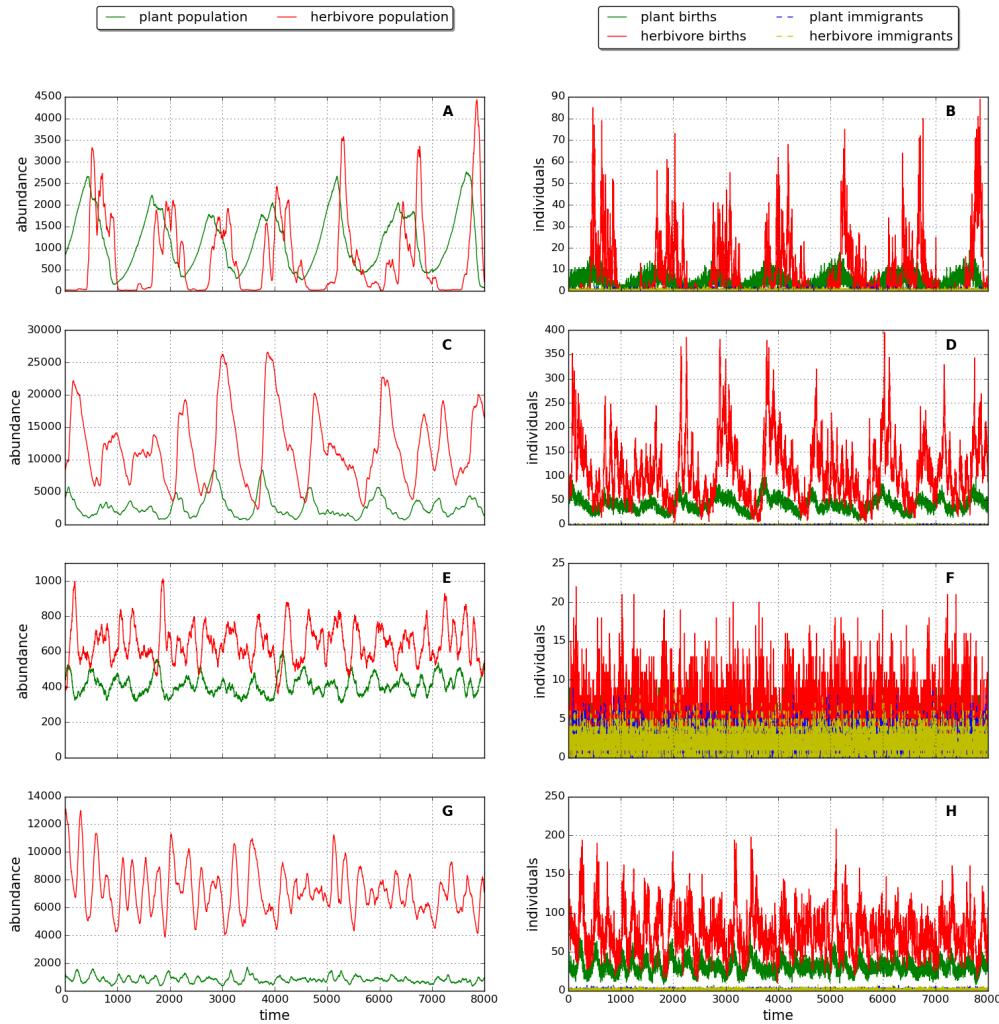
Figure 5.17 shows the dynamics of the IBM model, simulated with two species, for various parameter values. One species is a plant, the other is a herbivore which feeds on the plant. Throughout this section IBM simulations use a value of 1.0 for the parameter *HERB\_FRACTION*, rather than the default value of 0.7. This parameter defines the fraction of the energy/resource of a plant individual that is taken when it is fed on by a consumer. The use of the value 1.0 means that any feeding interaction results in the death of the plant individual, simplifying the interpretation of the modelling and the calculation of interaction strengths. Further to this parameter adjustment, we compare the effects of changing the immigration rate parameter (IR), and the reproduction rate parameter (RR), on two species dynamics.

The immigration parameter IR represents a source of noise in the dynamics. This fact, although intuitive, was revealed explicitly by *recurrence quantification analysis* in chapter 4. As we saw in section 5.3, the performance of the inference method can be highly sensitive to noise, especially when the functional response of the predator is non-linear (see below). However the noise introduced by immigration is slightly different from the noise modelled in section 5.3. In the ODE models multiplicative noise was used, such that the noise term vanished for zero populations and the *postulate of parenthood* was not violated (section 5.3.1). In the IBM immigration represents a source of individuals that is not dependent on species populations in the landscape, and therefore net immigration does not fall to zero when no individuals are

---

<sup>4</sup>make this explicit

#### 5.4. RESULTS: IBM AS DATA GENERATOR



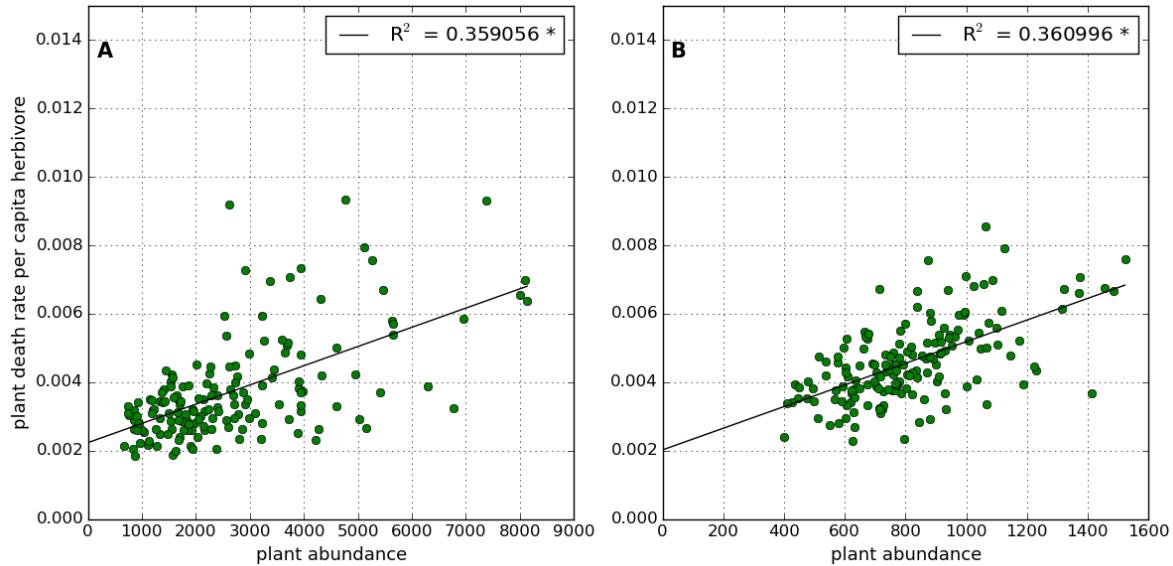
**Figure 5.17: Example dynamics of the IBM for two species** with different reproduction rates (RR) and immigration rates (IR). Here low and high RR are 0.01 and 0.1 respectively. Low and high IR are  $10^{-4}$  and  $10^{-5}$  respectively. **Left column:** Population dynamics of the two species. **Right column:** Time series of births and immigrations for both species. **First row:** low RR and low IR. **Second row:** high RR and low IR. **Third row:** low RR and high IR. **Fourth row:** high RR and high IR.

present. The IBM with non-zero IR violates the *postulate of parenthood*. However we argue that the noise introduced by immigration is still likely to cause systematic error in the estimates of interaction strengths. Importantly there is another source of noise resulting from the fact that the dynamics of the IBM are inherently stochastic. Individuals move around, interact and reproduce at random. It is not clear at this stage which source of noise is more significant. In an attempt to reduce the error introduced by immigration we use IR values that are low compared to those used in chapter ???. For *zero IR* we determined that the two species IBM dynamics was unstable. The herbivore species invariability went extinct (results not shown), which is consistent with results from chapter 4. Therefore two non-zero IRs were selected:  $10^{-4}$  and  $10^{-5}$ . The former value is equal to the lowest IR studied in chapter ???, but in the context of this chapter is referred to as the *high IR*. The latter value is an order of magnitude smaller, and therefore is referred to as the *low IR* in what follows.

The top two rows of figure 5.17 show dynamics simulated with the *low IR*, and the bottom two rows with the *high IR*. The left column shows the population dynamics, while the right column shows the number of immigrants and individuals born for both species on each time step. Panels A and E show dynamics that use the *default reproduction rate* ( $RR=0.01$ ). In the low IR case (panel A) we see that the populations display *relaxation-type* oscillations, with the herbivore population falling close to zero for different periods of time with intermittent spiking. Such oscillations are sensitive to noise because a small stochastic increase in population during the relaxation phase can induce spiking. It was experimentally determined that the presence of such oscillations hampered attempts to infer species interaction strengths (results not shown). By increasing RR the decline in the plant population, which causes the crash in predator population, is softened (a similar argument was used in section 4.2.2). Panels C and G show dynamics with an increased RR of 0.1. In the low IR case (panel C), although there are still large amplitude oscillations, the herbivore population no longer relaxes close to zero. In all simulations that follow we use the value  $RR=0.1$  in order to avoid relaxation-type oscillations.

Another consequence of increasing RR is that it increases the population levels of both species, on average. However, as we see from panels C and G, most of this benefit is conferred on the herbivore rather than the plant. Presumably the reason for this is strong predation by the herbivore, which regulates the plant populations. There is no predation pressure on the herbivore, only the effects of intrinsic mortality. The increased species abundances under high RR reduce dependence on immigration. In panels D and H the number if immigrants is barely visible, compared to the number of parented-births for both species. This observation suggests in this parameter regime that the contribution of noise resulting from immigration is low, compared to the inherent stochasticity of the IBM dynamics.

The dynamics in panels C and G in figure 5.17 correspond to the parameter values (low and high IR respectively) which we will use in section ?? to study two species systems. In later sections, where we study larger systems, the same parameter values are also used. Figure 5.2



**Figure 5.18: Functional response (FR) of the herbivore** at two different immigration rates (IR), experimentally derived from two IBM simulations. Both simulations are of two species plant-herbivore systems, with reproduction rate (RR) 0.1 and run for 10,000 time steps. Each green circle represents the number of plants consumed during a window of 50 time steps, divided by the mean herbivore abundance during that window, plotted against the mean plant abundance during the window. The black lines represent linear regression fits to the data.  $R^2$  values for the fits are given in legends, and significance at 95% confidence is indicated by \*. **Panel A:** Low immigration rate ( $IR = 10^{-5}$ ). **Panel B:** High immigration rate ( $IR = 10^{-4}$ ).

shows the functional response (FR) of the herbivore at both low and high IR values. The FR is the same concept as introduced in the context of ODE modelling (section 5.2.1). It defines the rate of prey consumption per-capita of predator. Here the FR is derived from the IBM simulation output by counting the number of plants consumed during a window of 50 time steps, and dividing that count by the mean herbivore abundance during the window. The plots in figure 5.2 show the resulting FR values over the course of two simulations (panel A: low IR, panel B: high IR). From these plots we see that the FR appears to be approximately linear in plant abundance, with some deviation from linearity resulting from noise. Linear regression fits to the FRs indicate slightly less deviation from linearity at the high IR value (panel B). This observation is counter-intuitive based on results from previous chapters, where we have seen immigration act as a source of randomness in community dynamics.

Figure 5.19 shows equivalent plots of the herbivore *numerical response* (NR). NR is a similar concept to FR, but defines the per-capita birth rate of a predator as a function of its prey population [127]. The NR is evaluated from the IBM simulation output in a manner analogous to that just explained for the evaluation of the FR. From the figure we observe that the NR is approximately linear. There is visibly more deviation from linearity in the low IR case (panel A),

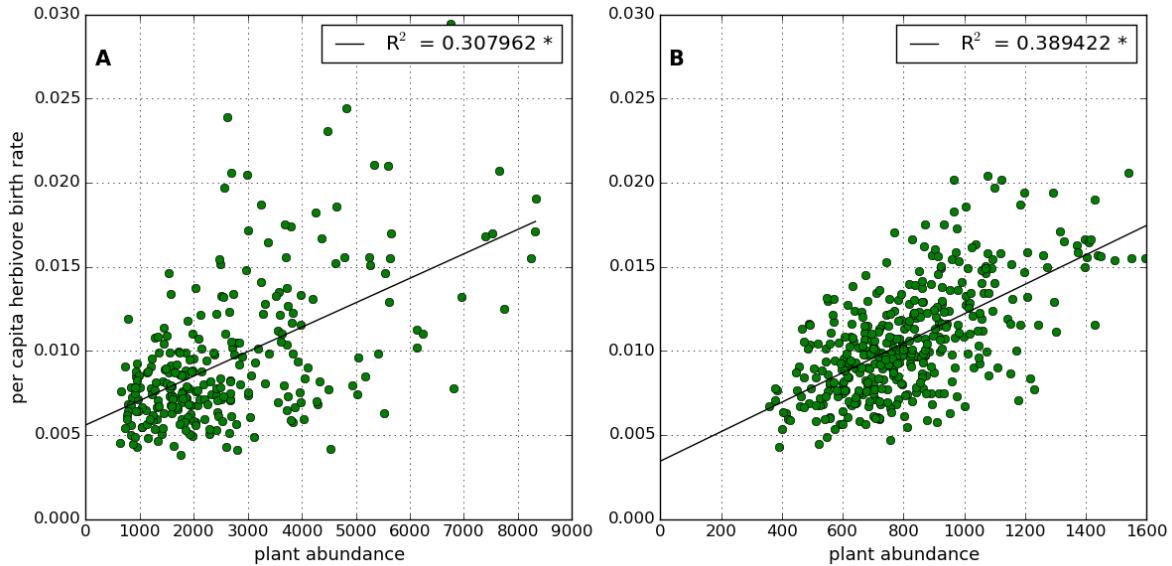
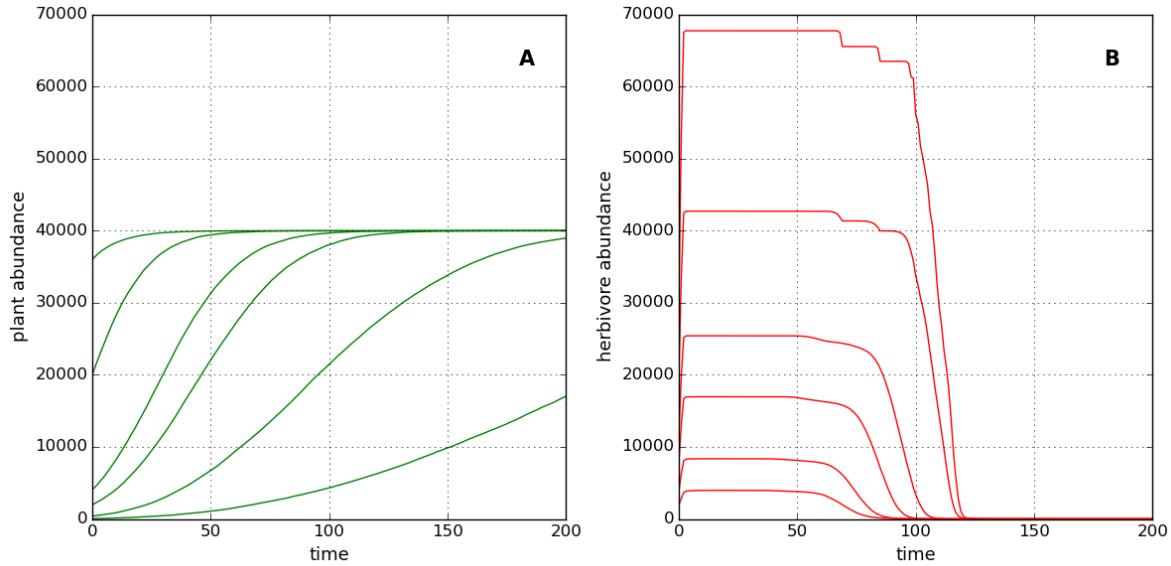


Figure 5.19: Similar to figure 5.18, but for the **numerical response (NR) of the herbivore**. Each green circle represents the number of herbivores born during a window of 50 time steps, divided by the mean herbivore abundance during that window, plotted against the mean plant abundance during the window.

than in the high IR case (panel B). This observation is confirmed by the  $R^2$  values of the linear regression fits. Therefore we conclude that noise resulting from stochasticity of the dynamics is more likely to be a source of error in the GLV fit, than is any non-linearity in the herbivore interaction functions. Indeed given that we do not model prey handling times in the IBM, it is reasonable that no signature of predator saturation is detected [60]. Furthermore it appears that stochastic effects are greater at high IR ( $10^{-4}$ ) than at low IR ( $10^{-5}$ ).

Figure 5.20 shows experimentally derived forms of the intrinsic growth and mortality functions of the plant (panel A) and herbivore (panel B). These results were obtained by simulating landscapes with different starting abundances of a single species, without any individuals of the other species. In all cases  $IR=0.0$ , such that the only new individuals are due to reproduction. The intrinsic growth of the plant (panel A) is the classic *logistic-type* sigmoidal shape that we expect when modelling a basal species. At low population densities the growth is near exponential, but at high population densities the growth rate is curtailed as the species approaches carrying capacity. In the case of plants the carrying capacity is equal to the number of landscape cells (40,000), since plant individuals may only occupy the *inhabitant space* in each cell (see section 2.3). The herbivore mortality function (panel B) is approximately sigmoidal. The initial conditions of the IBM are such that each individual begins with a random amount of energy. Therefore herbivores initially reproduce and the population level increases. However the population (for the experiments depicted) never reaches the herbivore carrying capacity, which is twice the number



**Figure 5.20: Intrinsic growth and mortality functions** derived experimentally from two species IBM simulations. **Panel A:** How the plant population grows from different initial abundances, in the absence of any herbivores. **Panel B:** How the herbivore population declines from different initial populations, in the absence of any plants.

of landscape cells (80,000) because non-basal individuals may occupy both the *inhabitant* and *visitor* spaces in a cell. The plateau in herbivore population is below carrying capacity because there is a finite amount of energy in the system. The decline in the population from the plateau is sharp, but display evidence of density-dependence similar to that observed for plant growth. The discrete steps visible at high population levels are possibly due to herbivores, which started with the same amount of energy, running out of energy at the same time. This feature suggests the possible existence of a *delay* in the herbivore response to change in the availability of food. However the main features of the intrinsic functions depicted, together with the near-linear functional forms seen in figures 5.2 and 5.19, suggest that an attempt to fit the GLV model to the IBM dynamics is valid.

### 5.4.3 Two species results

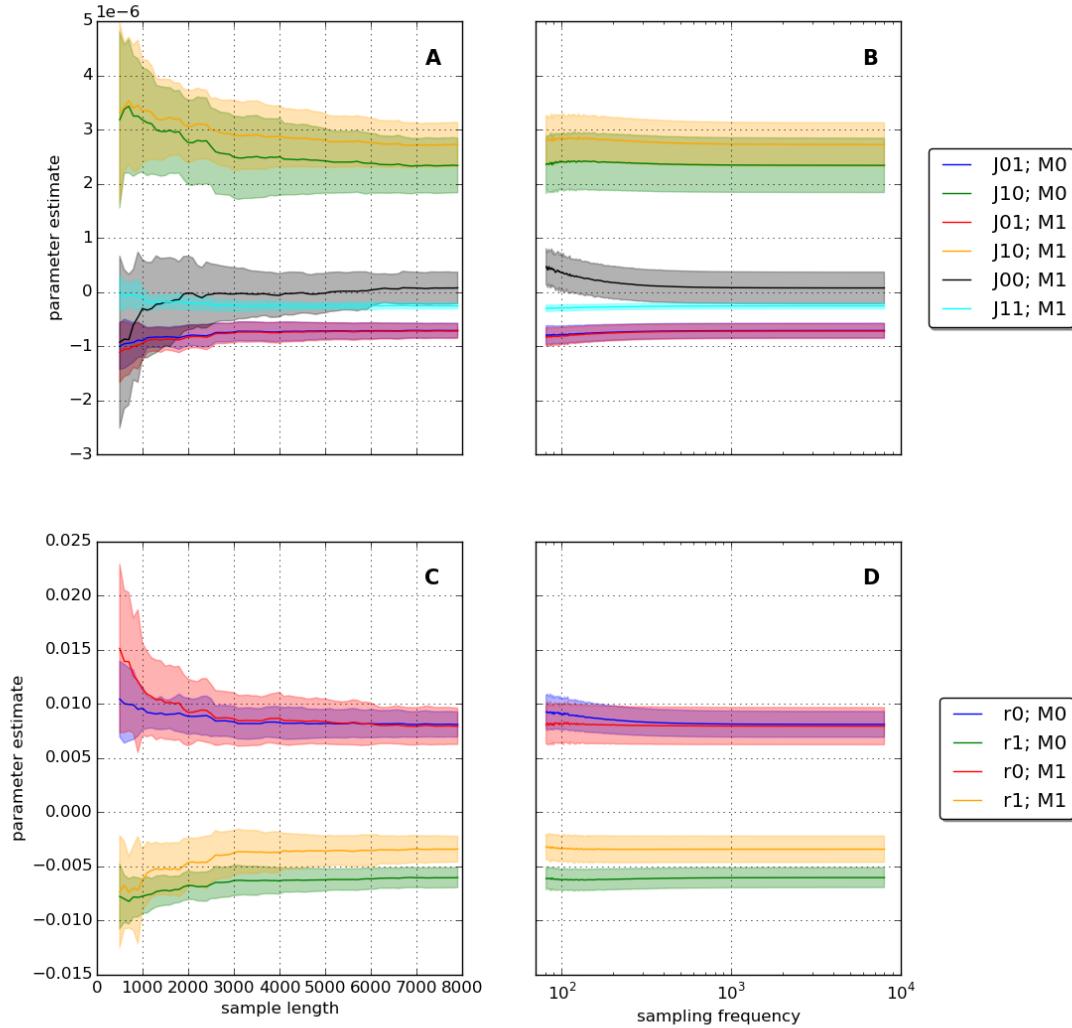
In this section results are presented for application of the inference method to *data streams* derived from two species IBM simulations. In all cases the simulations contain one plant and one predator species, and the parameters used are those described in the previous section. Therefore we compare the case of low IR ( $10^{-4}$ ) with high IR ( $10^{-5}$ ). We first demonstrate the convergence of the GLV parameter estimates under different sampling conditions, and then develop a method for quantifying the accuracy of the estimates. All simulations were run for 10,000 time steps, and the initial 2000 time steps were discarded to remove the initial transience (in accordance with

previous chapters).

Figure 5.21 shows how the GLV parameter estimates vary with sample length and sampling frequency in the case of *low IR*. As in section 5.3 sample length is varied by taking samples of increasing length from the beginning of the population dynamics, using all consecutive time points. Sampling frequency is defined as the total number of samples  $F$ , which are drawn at equal intervals from the full length of the dynamics. We compare the fits of two models,  $M_0$  and  $M_1$ .  $M_0$  is the GLV model with the constraint that intra-specific interactions are equal to zero (see section 5.4.1).  $M_1$  is the full GLV model without topological constraints. Each model is fitted to 25 replicate simulations, so we can compare the mean and variability in estimated parameter values. In general we observe that the parameter estimates converge as the both sample length and sampling frequency are increased. This convergence is more rapid for increasing sampling frequency, suggesting that it is better to use samples drawn from the full length of the dynamics. That is, for a given number of samples, low resolution sampling of the full dynamics performs better than high resolution sampling of a subset. We also observe that the introduction of intra-specific interactions ( $M_1$ ) does not significantly affect estimates of prey interaction strength ( $\hat{J}_{01}$ ), whereas it does affect estimates of predator interaction strength ( $\hat{J}_{10}$ ). The estimates of  $\hat{J}_{10}$  are also more variable than those of  $\hat{J}_{01}$ .

Figure 5.22 shows the equivalent convergence plots for the *high IR* case. Here the convergence of the estimates with sampling frequency is slower. From panels B and D we conclude that at least 1000 samples, distributed along the full length of the dynamics, are required to produce convergence of the estimates. As in the low IR case, we see that there is less variability in the estimates  $\hat{J}_{01}$ , than in  $\hat{J}_{10}$ , and that the latter estimates are more affected by the inclusion of intra-specific interactions. Note the scale on the y-axes of panels A and B are an order of magnitude larger than those in figure 5.21. The magnitudes of the estimates of inter-specific interaction strength ( $\hat{J}_{01}$  and  $\hat{J}_{10}$ ) are larger in the high IR case than in the low IR case. This observation is not consistent with results from chapter ??, where it was seen that reducing the IR increased the mean value of IS for the community as a whole. However the observation is consistent with the argument that reducing IR reduces the probability of interaction between individuals, because there are fewer individuals present in the landscape. We suggest that the disparity between these results and those of the previous chapter is due to the limit behaviour of the IS metric (as previously argued), which biases the community level result as some populations approach zero. However it may also be the case that the GLV parameter estimates converge on values which are incorrect (i.e. do not faithfully represent the interaction strengths of the system).

One possible way to determine the accuracy of the GLV parameter estimates is to simulate the GLV model using the fitted values, and compare the simulation to the original *data stream*. In figure 5.23 we use this method to compare the fits of  $M_0$  and  $M_1$  to a single *low IR* simulation. Both fits use a *sampling frequency* of 1000, based on the convergence results discussed above. In



**Figure 5.21: Convergence of parameter estimates for low immigration rate ( $IR = 10^{-5}$ ).** Solid lines represent mean values, shaded areas represent  $\pm 1$  standard deviation, over 25 replicates. M0 indicates GLV model fit without intra-specific interactions ( $J_{ii} = 0$ ). M1 indicates GLV model fit without constraints on topology (see text). **Top row:** interaction strength estimates. **Bottom row:** growth rate estimates. **Left column:** convergence as sample length increased. **Right column:** convergence as sampling frequency increased.

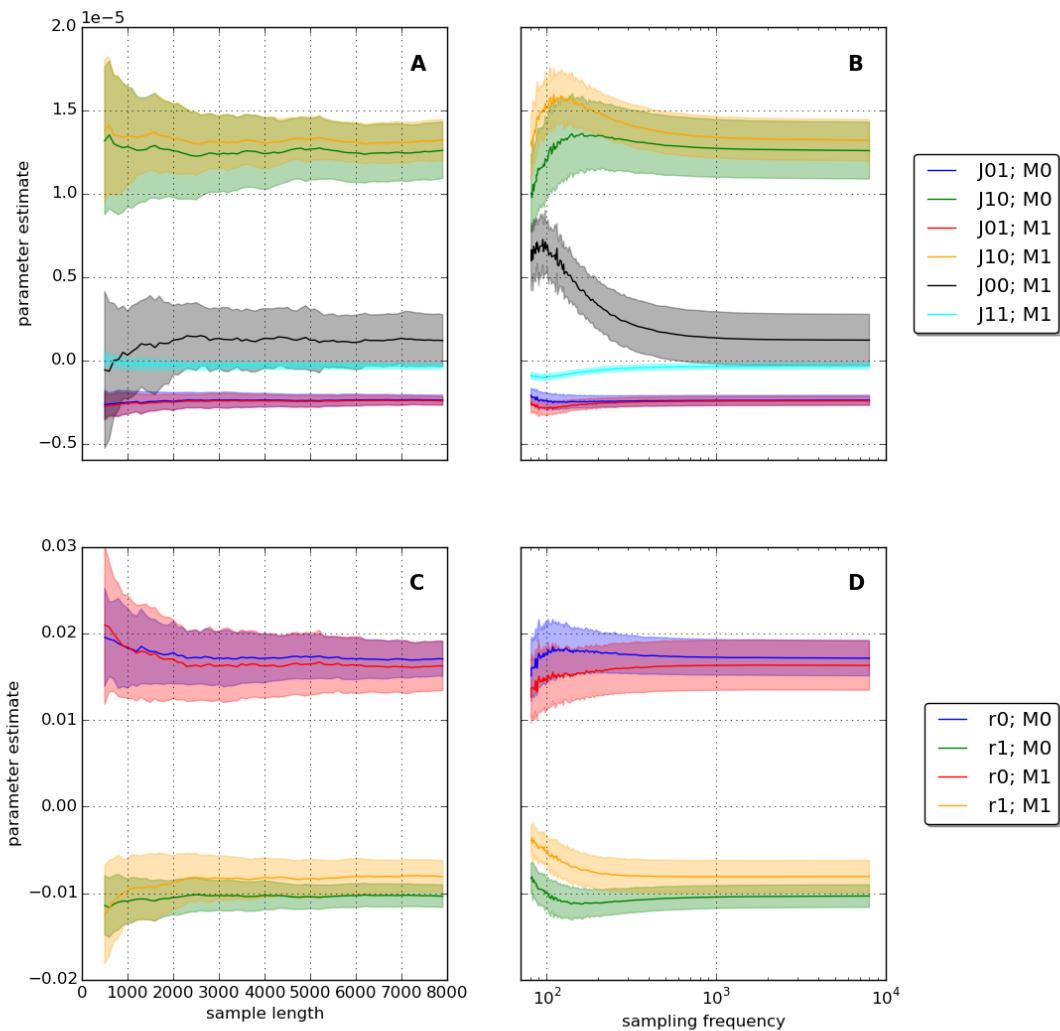
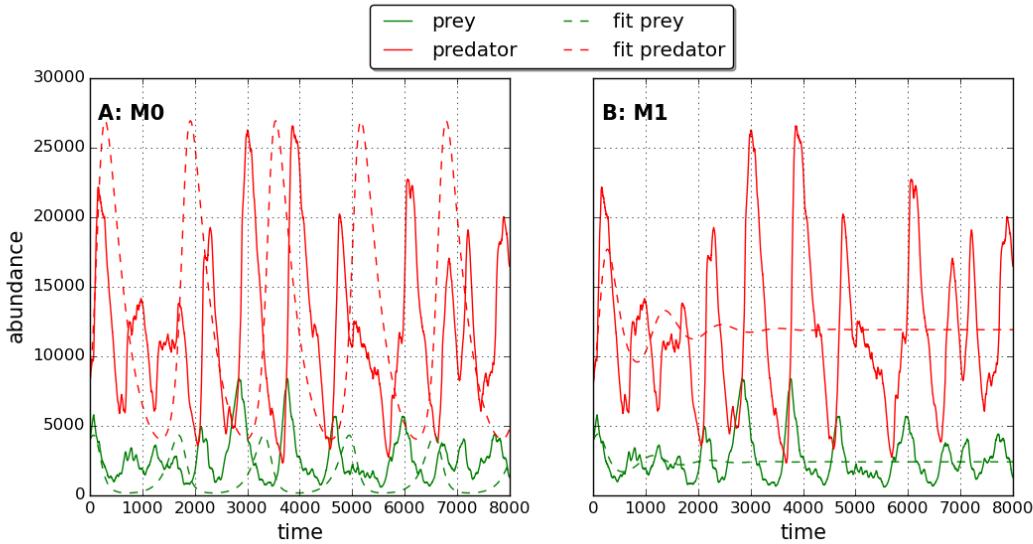


Figure 5.22: Similar to figure 5.21, but for **high immigration rate** ( $IR = 10^{-4}$ ).



**Figure 5.23: Comparison between true and fitted dynamics** for two different fitted models: M0 and M1. Solid lines represents the simulated IBM dynamics at **low IR** ( $10^{-5}$ ). Dashed lines indicate a deterministic simulation of the fitted GLV, from the same initial conditions as the IBM time series. **Panel A:** M0 is GLV model with constraint that  $J_{ii} = 0$  (no intra-specific interactions). **Panel B:** M1 is full GLV with no constraints.

panel A we see that the simulated M0 fit produces oscillatory dynamics of comparable amplitude and period to those in the *data stream*. In panel B we see that the introduction of intra-specific interactions dampens the oscillations in the fitted model (M1). However stochastic forcing may induce oscillations in a system with a stable equilibrium, such as this M1 fit. Therefore, under the correct noise conditions, the M1 fit may produce dynamics the are similar to the *data stream*. Furthermore the shapes of the intrinsic functions, seen in section 5.4.2, suggest the presence of intra-specific effects in the IBM. Therefore we conclude that simulating the fitted GLV is not an appropriate way to determine to accuracy of the fit.

Rather than simulating the fitted GLV model, we assess the predictions that the fitted model makes about the birth and death rates of both species. To compute these predictions the relevant parameters of the GLV are multiplied by some combination of the species population vectors. For example the prediction for *predator births* is given by  $\hat{J}_{10}X_0X_1$ , where  $X_0$  and  $X_1$  are the true populations of the two species from the IBM simulation. The full set of definitions for the predicted rates of the fitted models are given in table 5.2. The GLV model fit without intra-specific interactions is named M0, as before. The GLV fit with intra-specific interactions raises the question of whether intra-specific interactions contribute to the birth or death rate of a species. Therefore the models labelled M1 and M2 refer to the same GLV fit, but M1 has the intra-specific terms added to the birth rate predictions, where M2 has them added to the death rate predictions. All rate predictions are plotted in figure 5.24 for a low IR simulation, and in figure 5.25 for a high

IR simulation. The model fits in all cases us a sampling frequency of 1000. The true birth and death rates for both species are plotted in green, for visual comparison with the predictions.

model	M0	M1	M2
prey births	$\hat{r}_0 X_0$	$\hat{r}_0 X_0 + \hat{J}_{00} X_0^2$	$\hat{r}_0 X_0$
predator births	$\hat{J}_{10} X_0 X_1$	$\hat{J}_{10} X_0 X_1 + \hat{J}_{11} X_1^2$	$\hat{J}_{10} X_0 X_1$
prey deaths	$\hat{J}_{01} X_0 X_1$	$\hat{J}_{01} X_0 X_1$	$\hat{J}_{01} X_0 X_1 + \hat{J}_{00} X_0^2$
predator deaths	$\hat{r}_1 X_1$	$\hat{r}_1 X_1$	$\hat{r}_1 X_1 + \hat{J}_{11} X_1^2$

Table 5.2: The way in which the various rate estimates are calculated from the fitted model parameters, based on the observed populations. The vectors  $X_0$  and  $X_1$  are the full abundance time series of the plant and herbivore species respectively. Multiplication of these vectors is *element-wise*.

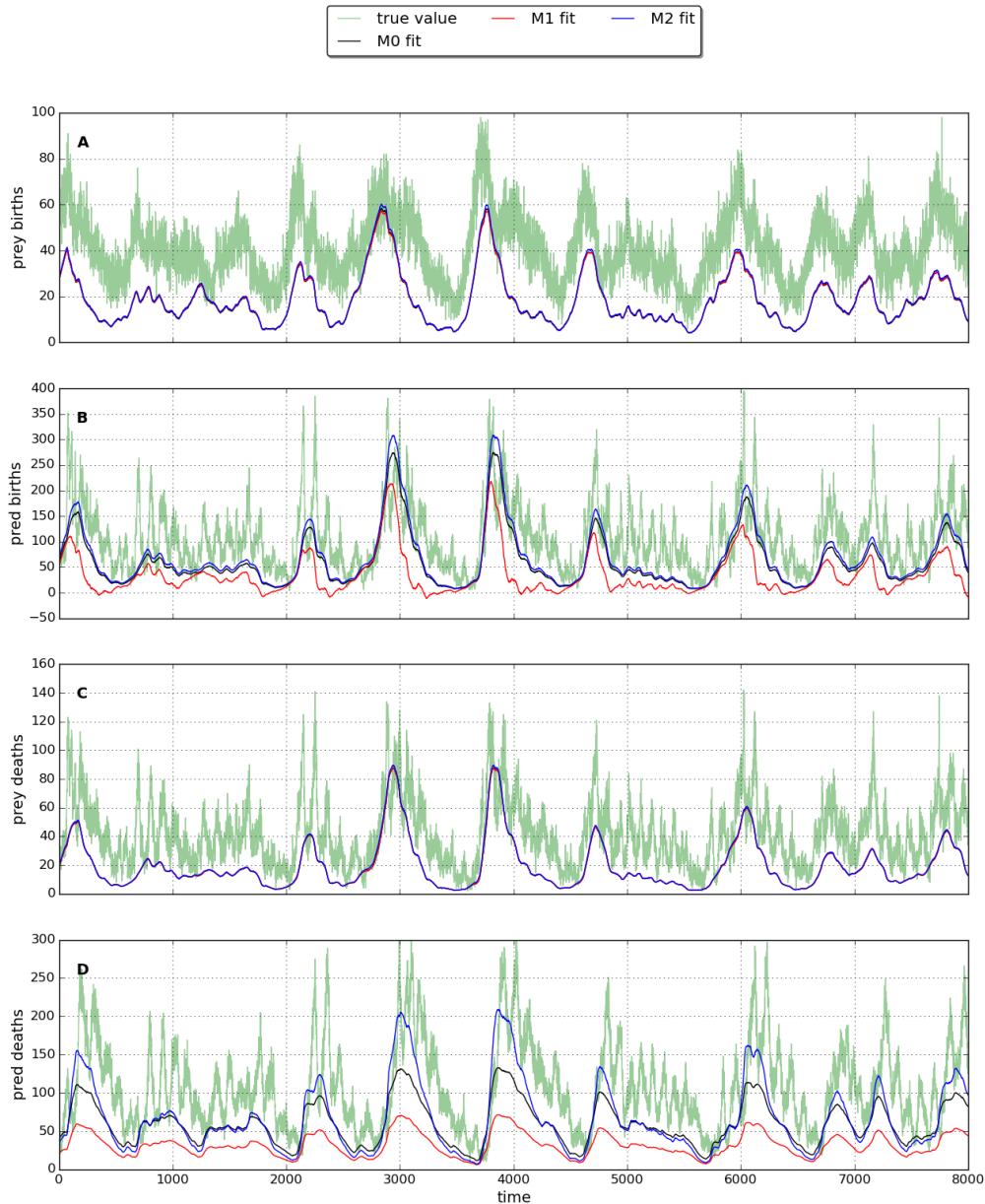
From figures 5.24 and 5.25 we make the following observations. The periods of oscillation in the predicted rates match those of the true values closely. This is no surprise given that the predictions are calculated using the true population time series. However, in most case the predictions are also of comparable magnitude to the true values. Therefore the GLV fit can produce reasonable predictions of all demographic rates at any time point, given the abundance of both species. In the predictions of prey births and deaths there is little different between models M0, M1 and M2 (panels A and C, both figures). In the predictions of predator births and deaths, M1 clearly gives inferior results at low IR (panels B and D, figure 5.24). At high IR a difference is visible between the estimates of the three models (panels B and D, figure 5.25) but it is hard to tell which model produces the best predictions. In order to quantify the performance of the different models in predicting species births and deaths we define a *relative error metric*

$$(5.41) \quad \langle RE \rangle = \frac{1}{T} \sum_{t=1}^T \left( \frac{|b(t) - p(t)|}{|b(t)|} \right),$$

where  $T$  is the number of time steps in the dynamics,  $b(t)$  is the true number of births or deaths on time step  $t$ , and  $p(t)$  is the number of births or deaths predicted by the fitted model. This metric quantifies the mean relative deviation of the predictions from the true values over the full time series.

Figure 5.26 summarises the relative error in the predicted births and deaths over an ensemble of 25 replicate simulations at low IR. The models are fitted to the simulated dynamics using a sampling frequency of 1000, as before. The fitted model parameters are then used to predict the

#### 5.4. RESULTS: IBM AS DATA GENERATOR



**Figure 5.24: Births and deaths predicted** by three different model fits: M0, M1, M2 (see text for model definitions). Green line is the true number of births and deaths for each species on each time step, recorded from the IBM simulation. Coloured lines represent model predictions of the number of births and deaths, based on the true population level of both species at each time step (see text for further details). The simulation was run using **low IR** ( $10^{-5}$ ).

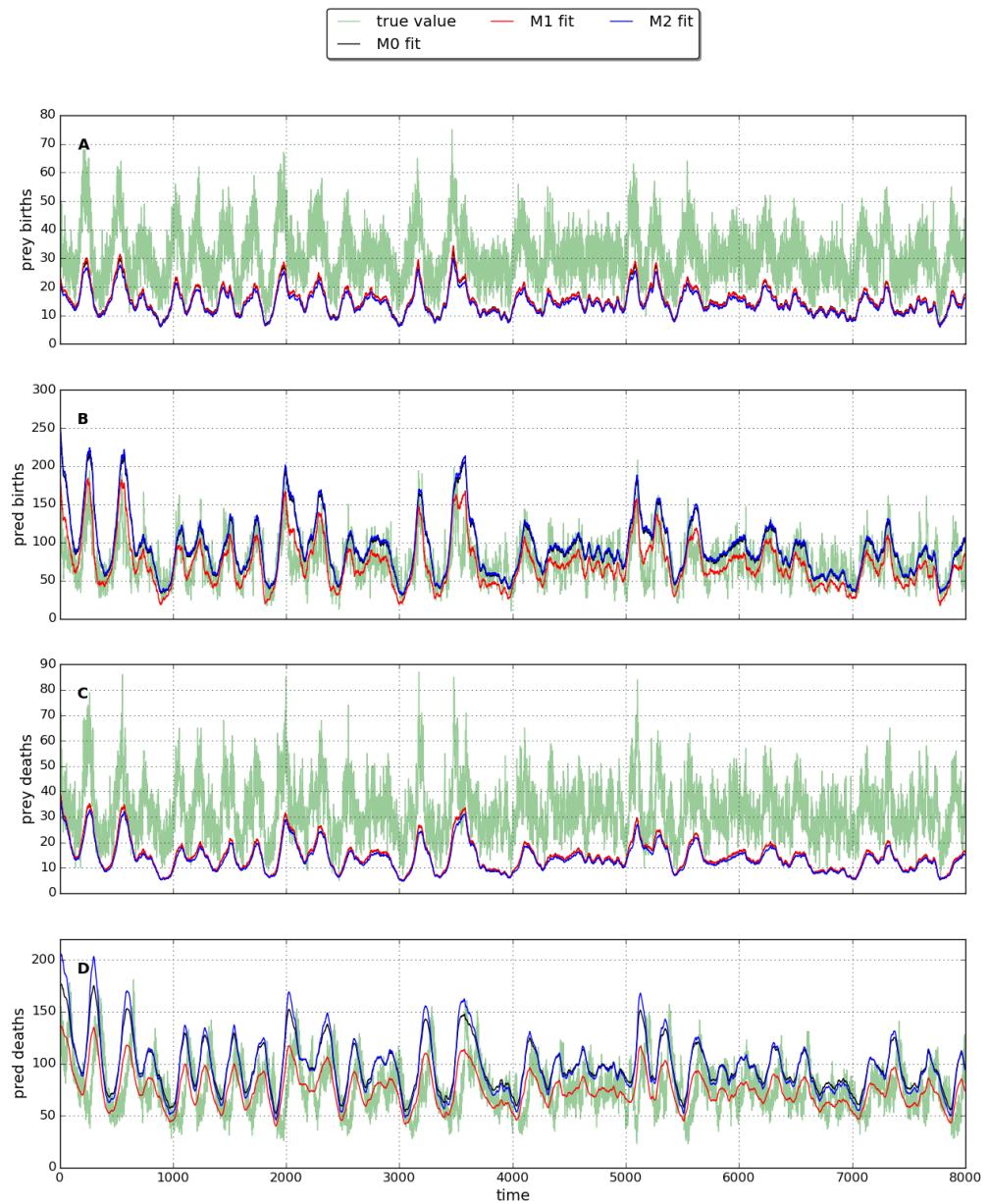


Figure 5.25: Similar to figure 5.24, but for an IBM simulation at **high IR** ( $10^{-4}$ ).

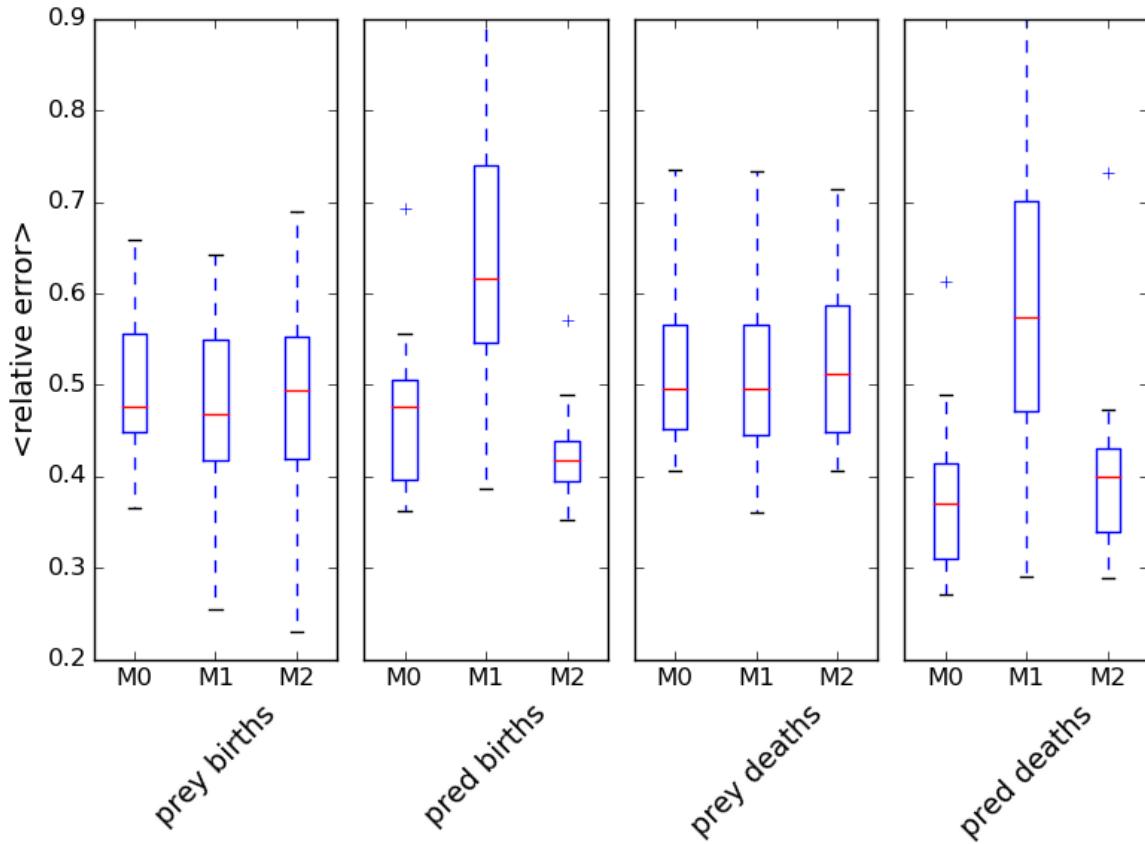


Figure 5.26: **Relative error in the rates estimates** displayed in figure 5.24 (prey and predator births/deaths). The relative error metric is defined in equation (5.41). The relative error statistics illustrated are derived from model fits to 25 replicate IBM simulations at **low IR** ( $10^{-5}$ ). Boxes extend to the first and third quartiles of the data, and the red line indicates the median. Whiskers extend to show the range of the data, up to a maximum of 1.5 times the inter-quartile distance beyond which points are shown as outliers.

births and deaths of both species, and the relative errors (RE) calculated as just described. For prey births and deaths there is little difference between the three models. Based on the median RE (red lines), the model M1 produces the best estimates by a narrow margin. For the predator estimates M1 clearly has higher errors, while M2 and M0 perform best for births and deaths respectively. Figure 5.27 shows the equivalent results for 25 simulations at high IR. The prey estimates have similar, or slightly higher, errors than in the low IR case, whereas the predator estimates show significantly lower errors<sup>5</sup>. In the high IR case M2 performs the worst for the prey estimates, and M1 performs the worst for the predator estimates. Comparing across both low and high IR it appears that M1 is the correct model for the prey, while M2 is the correct model for the predator. However the inclusion of intra-specific interactions does not necessarily

<sup>5</sup>This seems an important feature! - comes up again later. Discussion.

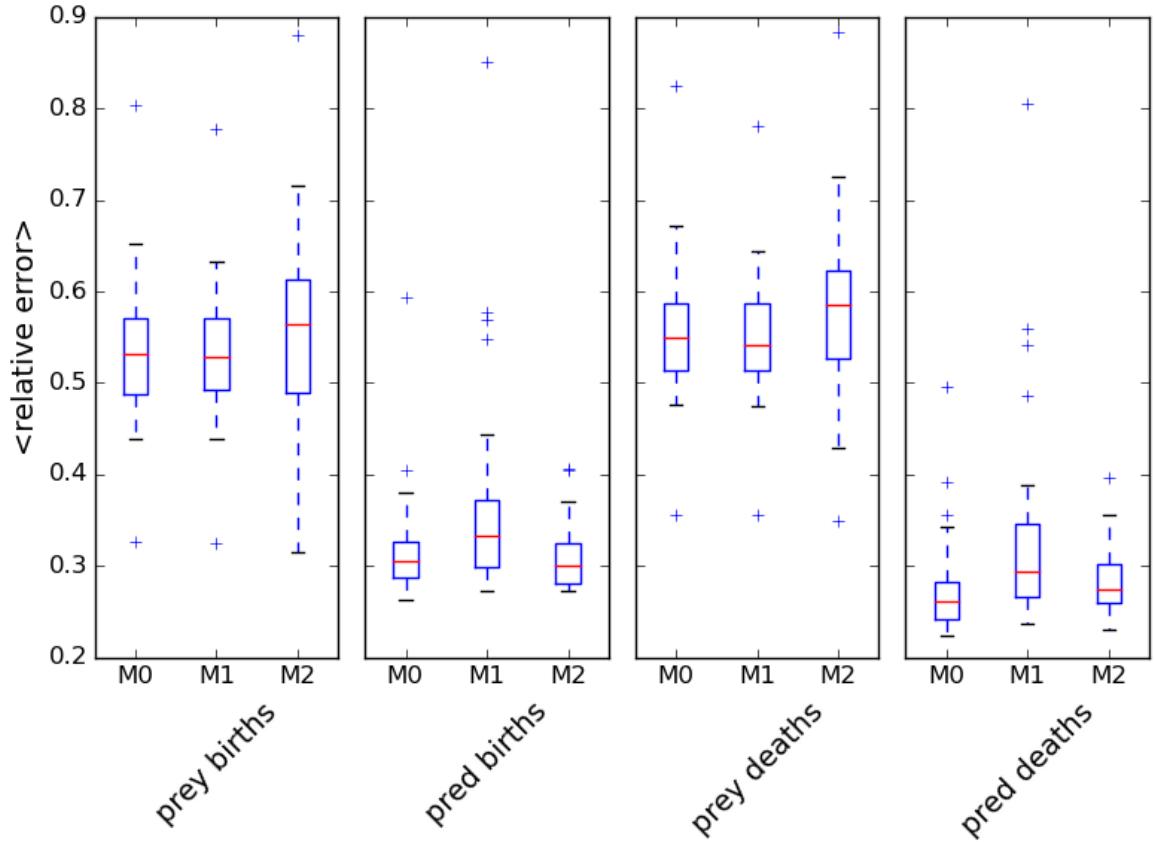


Figure 5.27: Similar to figure 5.26, but for results of model fits to 25 replicate IBM simulations at **high IR** ( $10^{-4}$ ).

improve the rate predictions over those of M0. The use of M1 for prey predictions, and M2 for predator predictions, is consistent with the intrinsic functions derived in section 5.4.2. That is, intra-specific interactions contribute to prey growth rates, but to predator mortality rates.

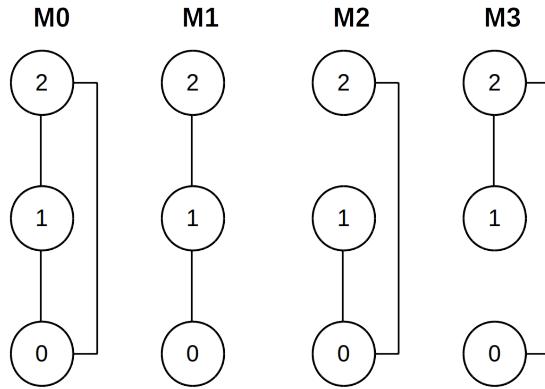
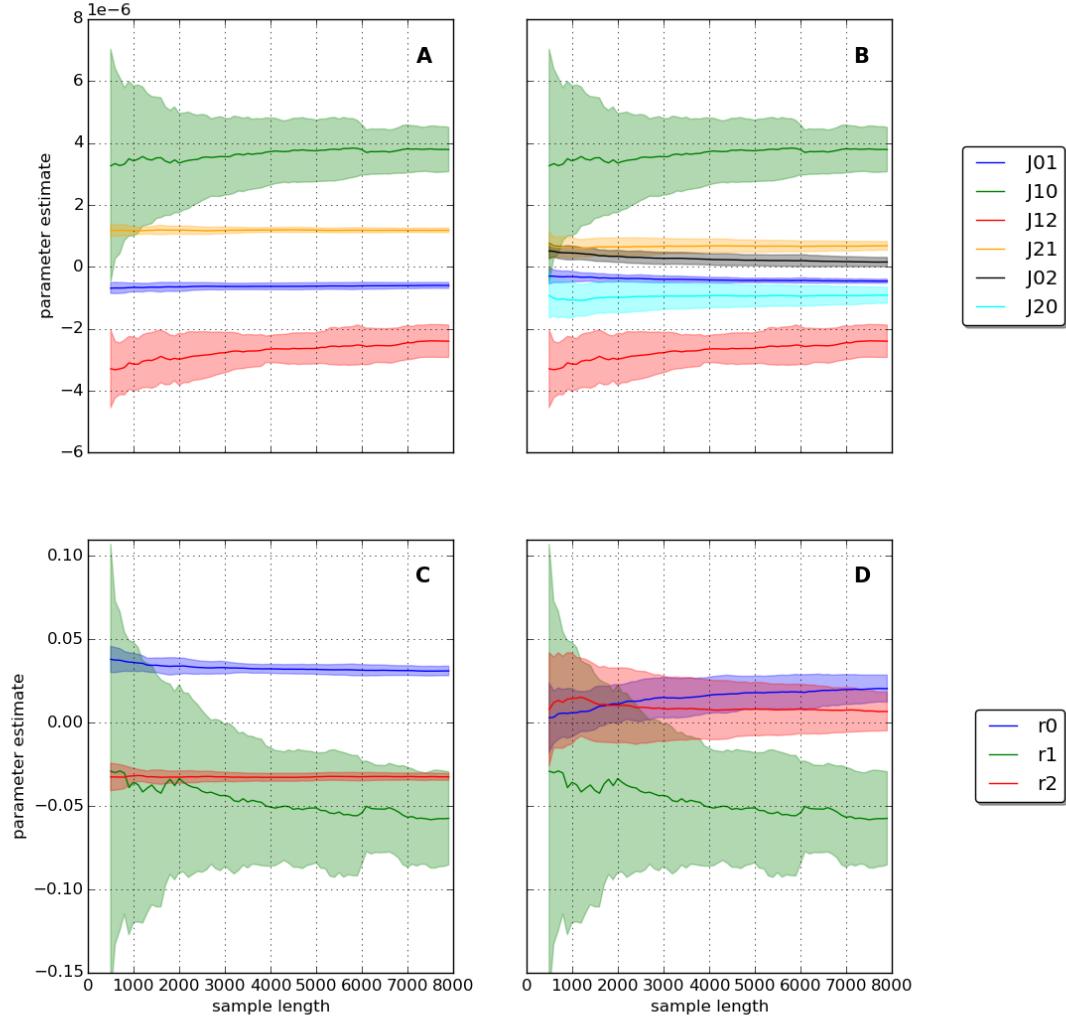


Figure 5.28: The four candidate models investigated in this section. M0: fully connected. M1, M2, M3: one links removed. M1 is the ‘correct’ model i.e. the food chain simulated by the IBM, to which the models are fitted. All four models include intra-specific interactions (self-loops), although these are not drawn.

#### 5.4.4 Three species results

In this section results are presented for fitting to 3 species IBM simulations of a *food chain*. Each simulation contains a single plant species, a single herbivore species, and a single predator species which feeds only on the herbivore. This scenario represents the first test of the inference method in correctly identifying the interaction network topology. There is a potential link, between the plant and predator species, that is not present in the underlying interaction network. We investigate whether the method can detect the absence of this link. In all cases the GLV is fitted *with* intra-specific interactions. According to the analysis in the previous section such interactions are taken to contribute to plant births, and to the deaths of the other two (non-basal) species, when rate predictions are made. We investigate the fitting of four candidate models (M0-3) with different constraints on interaction topology. These topologies are shown in figure 5.28, and are the only four possibilities that are connected (i.e. no isolated species). All models are subsets of M0, with M1 representing the true food chain that is simulated.

Figure 5.29 shows the convergence of parameter estimates with sample length for models M0 (right column) and M1 (left column), at high IR. The results at low IR are similar, and therefore not shown. The convergence depicted here is less convincing than in the two species case. In particular the estimates  $\hat{J}_{10}$ ,  $\hat{J}_{12}$  and all  $r_i$ , do not settle down as sample length is increased. A further concern is that, for all sample lengths,  $\hat{J}_{02} > 0$  and  $\hat{J}_{20} < 0$ . These parameters correspond to the spurious link (not present in the true network), and the signs quoted are indicative of the *plant species consuming the predator species*. It is worth noting that, when fitting M0, parameter that correspond to missing links will never be estimated as *exactly* equal to zero. The hope is that either the model fit reveals that such parameters are of small magnitude, or that there is some other means by which to determine to correct topology.



**Figure 5.29: Convergence of parameter estimates with sample length for high immigration rate ( $IR = 10^{-4}$ ).** Two different models fitted to three species food chain dynamics. Inter-specific interactions (top row), and intrinsic parameters only (bottom row). Solid lines represent mean values, shaded areas represent  $\pm 1$  standard deviation, over 25 replicates. Panels A and C show fits of model M1. Panels B and D show fits of the model M0. These models are illustrated in figure 5.28. Intra-specific interactions are included in the fits, but are not plotted for simplicity.

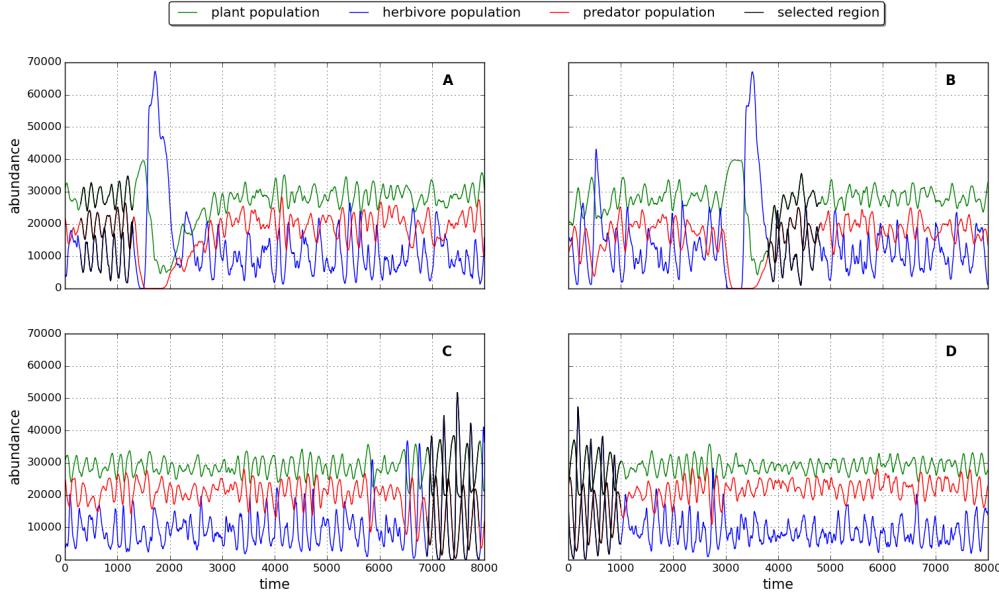


Figure 5.30: **Example dynamics of 3 species IBM** food chain, four different simulations (panels A,B,C,D). Top row: low IR ( $10^{-5}$ ). Bottom row: high IR ( $10^{-4}$ ). Regions plotted in black are those selected as *data streams* for GLV fit (see text for details of how selected).

Looking at the simulated food chain dynamics in figure 5.30, we speculate that the lack of convergence, and possibly the spurious interaction, may result from features of the dynamics. The dynamics shown are highly variable and appear *quasi-periodic*. It may be that the GLV struggles to fit to large regions of such dynamics. In an attempt to overcome this problem we employ a method to select the optimum region of the dynamics for fitting the GLV. The selection method consists of scanning a sampling window of length 1000 time steps along the full dynamics. The model M0 is fitted to the dynamics within the window and the error functions (equation (5.34)) of the three species are summed. We then select the window with the minimum total error in the fit. These selected regions are plotted in black for the four simulations shown in figure 5.30. The regions appear to share the common feature that they display oscillations of a relatively constant amplitude and period, compared to the rest of the dynamics. This observation supports our belief that the complexity of the dynamics negatively affects the GLV fit. In the analysis that follows all models are fitted to regions of the dynamics selected in the way just described.

Figures 5.31 and 5.32 show birth/death predictions of the fitted models to a single three species simulation at low IR and high IR respectively. Here we see significant difference between the quality of the predictions made by the different models. Note that for each species there is always a model (M1-3) that is equivalent to model M0 for that species. This is because in each of the models M1-3 one species is allowed to interact with both of the others, as they all are in model M0 (see figure 5.28). For most of the predictions at both IR values M3 (yellow) visibly performs

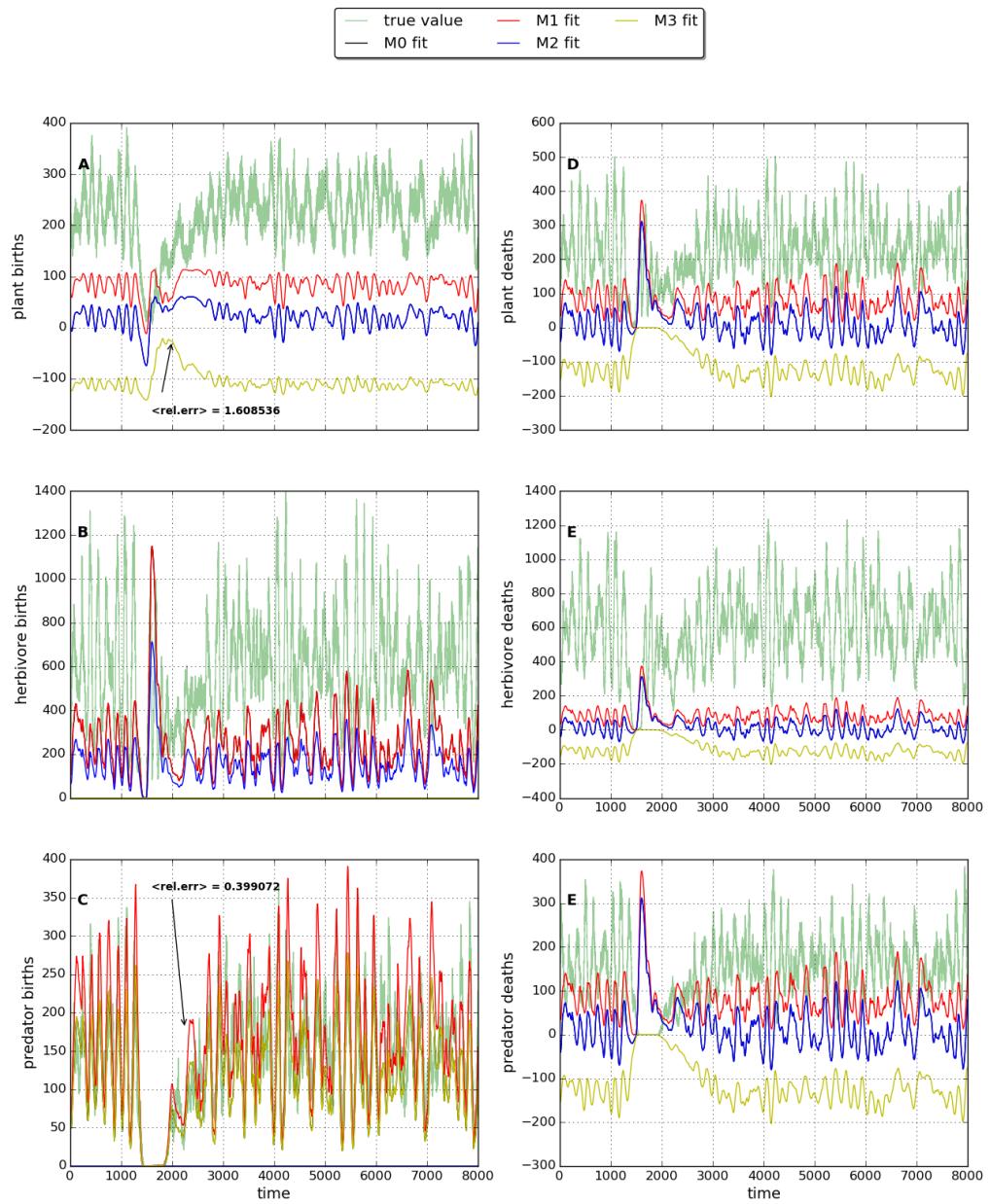


Figure 5.31: Similar to figure 5.24, but with different models fitted to a three species simulation at **low IR**. The simulation is the same as that in panel A of figure 5.30. The models M0-3 are depicted in figure 5.28. Arrows indicate the value of the relative error metric (RE) for two example model predictions.

#### 5.4. RESULTS: IBM AS DATA GENERATOR

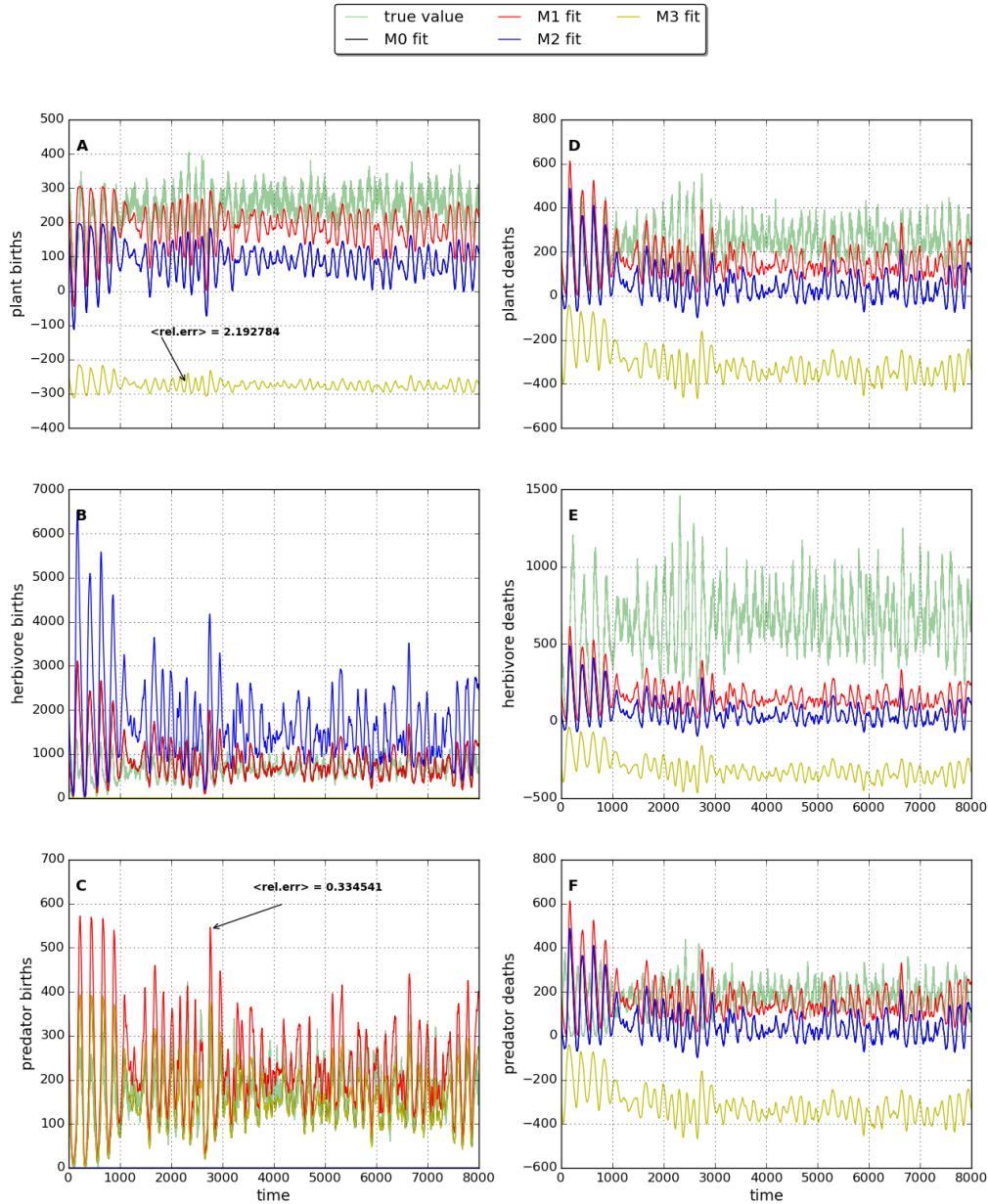


Figure 5.32: Similar to figure 5.31, but for **high IR**. The simulation is the same as that in panel C of figure 5.30.

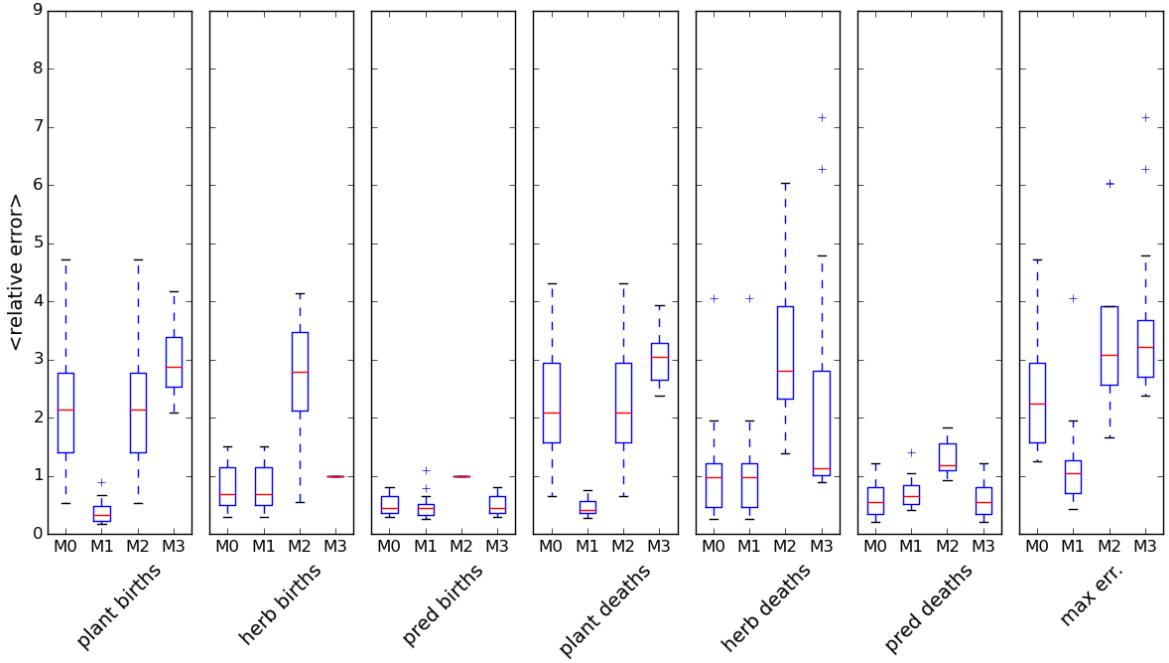


Figure 5.33: Similar to figure 5.26, but with different models fitted to 25 three species simulations at **low IR**. This figure includes an additional statistic labelled *max err.*, which is the maximum relative error in any prediction for a given simulation.

the worst. This is the model with the link between plant and herbivore removed. Also M2 appears to perform worse than M1 in most cases. M2 is the model with the link between herbivore and predator removed. It is reassuring that the removal of links which are present in the true network results in predictions that are worse than those of M1. To confirm the repeatability of this observation we employ the relative error metric (5.41) from the previous section, the determine the accuracy of model predictions over ensembles of 25 replicate simulations.

Figures 5.33 and 5.34 show the relative errors in birth/death predictions for the ensembles of simulations at low and high IR respectively. We see that the predictions of M1 perform better than those of M2 and M3 in almost all cases (according to the median and the range of RE). Furthermore M1 performs at least as well as M0, the unconstrained GLV fit, is most cases. The exceptions to these observations are in the predator deaths at low IR, and both predator births and deaths at high IR. In these cases models M0 and M3 perform at either better than, or equally as well as M1. We conclude that the inclusion of the spurious link between plant and predator species improves the prediction of the demographic rates of the predator by the fitted model. We also reiterate that when this link is included the parameter estimates  $\hat{J}_{02}$  and  $\hat{J}_{20}$  suggest that the predator is being eaten by the plant (see for example figure 5.29). We return to this strange feature of the results in section 5.4.6. In the final panel of figures 5.33 and 5.34 we include the statistic *maximum error*, which is the largest relative error out of all rate predictions for a given

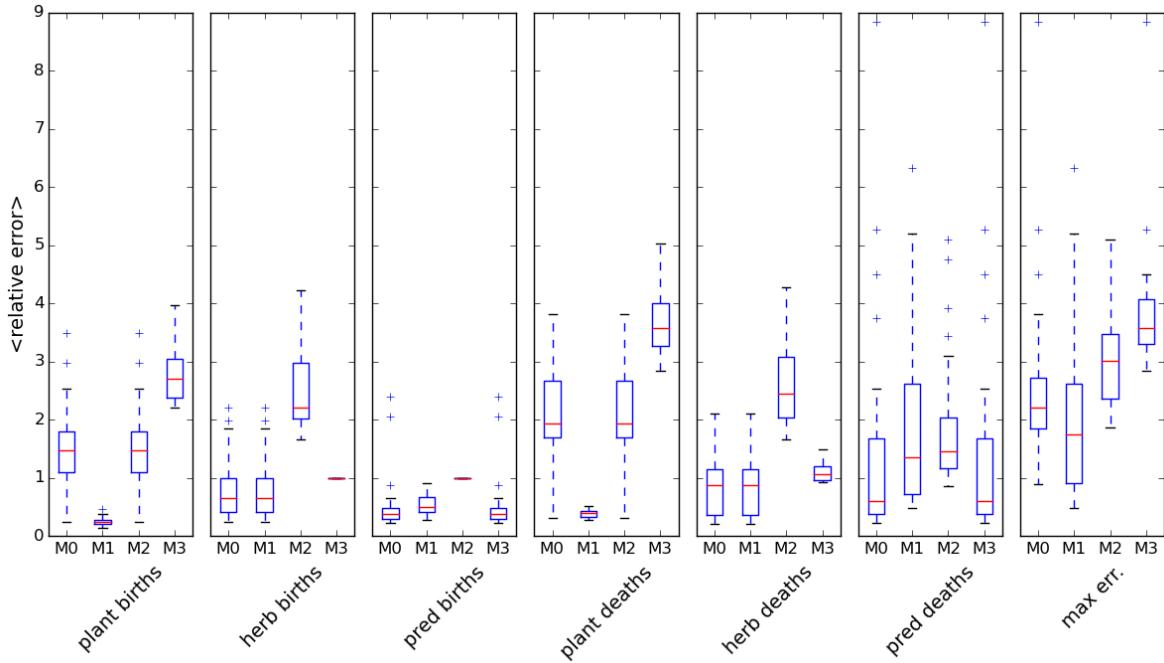
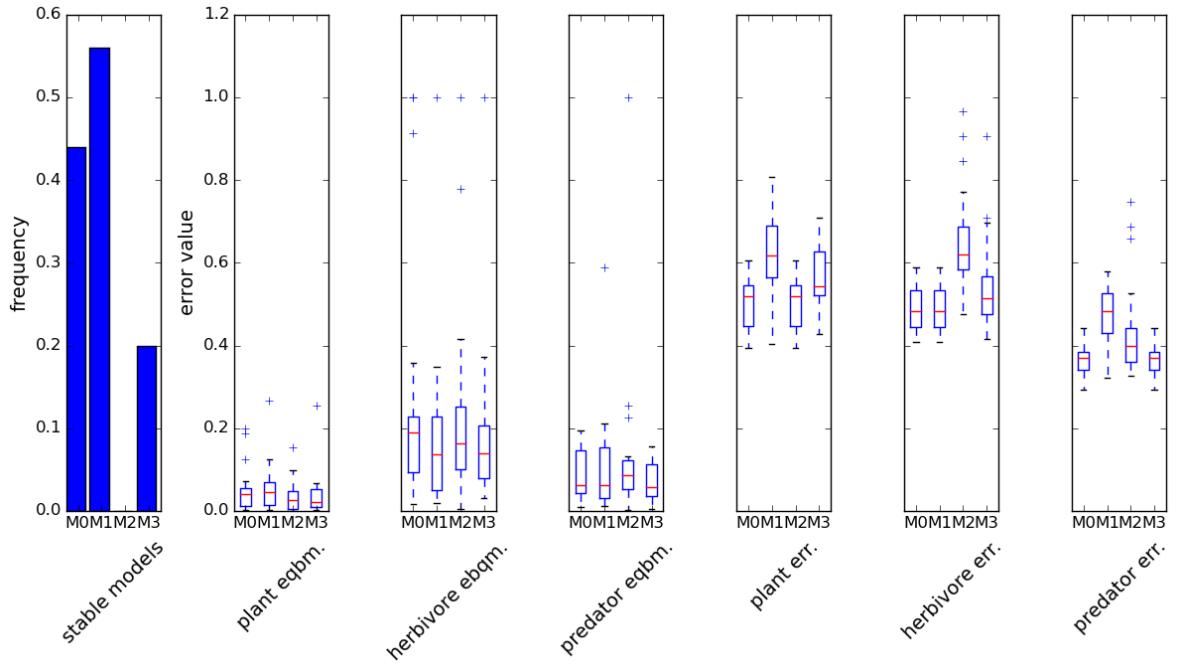


Figure 5.34: Similar to figure 5.33, but for 25 simulations at **high IR**.

simulation. For example, in figure 5.32, the maximum relative error for M3 is in the prediction of plant births. The error takes a value of 2.19, which is indicated by the arrow. The *maximum error* statistic allows us to determine which model fit provides the best predictions across all demographic rates. In both the low and high IR cases M1, the model with the correct topology, performs the best.

The correct model (M1) may produce the best rate predictions, but without knowledge of the true rates we cannot use this fact to identify the correct topology. We now attempt to identify which model (M0-4) represents the correct topology using only knowledge of the fitted models, and information about the IBM population dynamics (i.e. no knowledge of true demographic rates). We conduct three checks on the fitted models. First we evaluate the Jacobian of the model and determine if the equilibrium is locally stable (see section 5.2.1). Secondly we calculate the relative error between the model equilibrium and the long term average of the population dynamics for each species. Thirdly we compute the error function of the model fit  $E(\hat{J}_i)$  for each species  $i$  (equation (5.34)), normalised by the sum of the first derivatives of the population dynamics of species  $i$ <sup>6</sup>. The results of these three checks are summarised in figures 5.35 and 5.36 for the low and high IR ensembles respectively. At low IR, M1 does not perform consistently better than the competing models in terms of either equilibrium or error function values. At high IR, M1 performs best in terms of the predicted equilibrium values for all species, although the predator

<sup>6</sup>This explanation is not clear, and does not justify the use of this measure



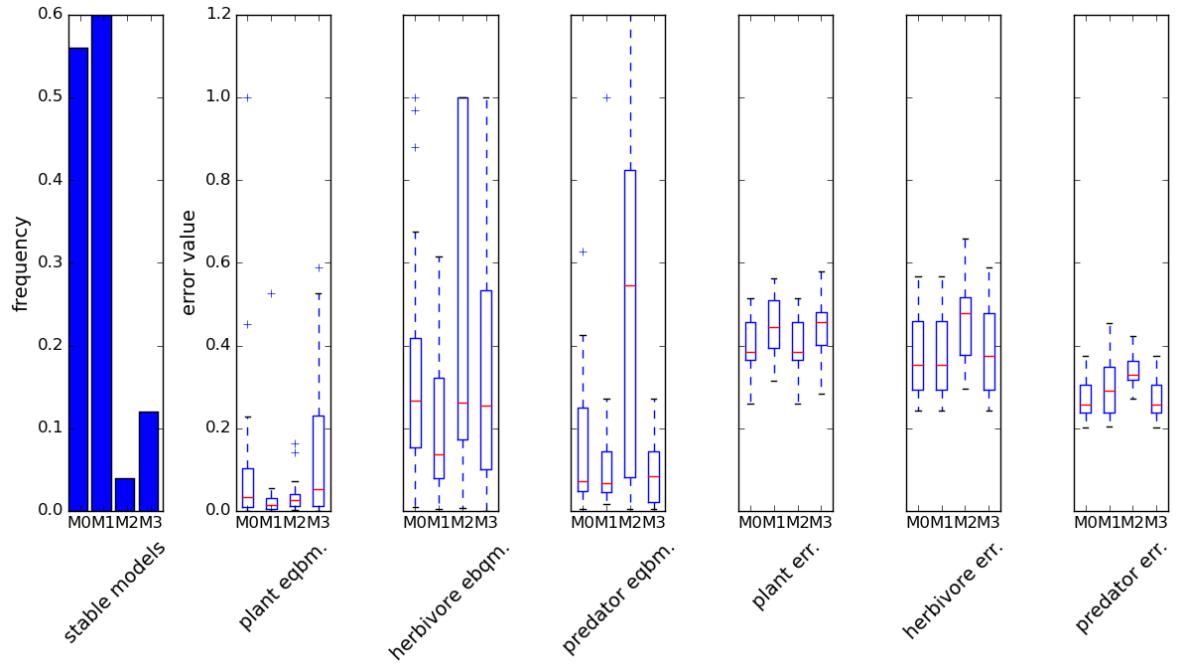
**Figure 5.35: Evaluating the performance of the different model fits** over 25 replicate simulations at **low IR**. Stability (panel 1) is determined from the Jacobian of the fitted model (see text). The errors in the equilibrium of each species (panels 2-4) represent the relative error between the equilibrium population of the fitted model and the long term average of the species population dynamics. The final three errors (panels 5-7) are the normalised error function of the model fit for each species (see text).

equilibrium results are comparable to those of M3. However M1 again does not display the lowest error function value for any species. In both the low and high IR cases the fitted M1 model is stable more frequently than any other. M0 is also stable with relatively high frequency, compared to M2 and M3 which are usually unstable. If presented with a single *data stream* of three species from which to infer the correct interaction topology, it may be possible to do so with reasonable confidence based on these results (with particular focus on stability). However no conclusive method for doing so emerges from the three checks presented. We concluded that the identification of the correct topology from three species food chain simulations remains an open problem<sup>7</sup>.

#### 5.4.5 Five species results

In this section we briefly consider the application of the inference method to a five species system. The step from three to five species significantly increases the complexity of the problem.

<sup>7</sup>Need to comment on difference between low and high IR? Esp. equilibrium values. And the sustained success of M3?

Figure 5.36: Similar to figure 5.35, but for **high IR**.

Given the difficulty of identifying the correct interaction topology for three species, we that anticipate this goal may be impossible using the current methodology. However, based on the three species results, we may expect that fitting the correct model (again called M1) will provide the best predictions of demographic rates. With five species there are a total of ten possible inter-specific interactions. This fully connected topology is shown as M0 in figure 5.37. There are a combinatorially large number of subsets of M0, depending on how many links L are included. For a network with L links there are  $\binom{10}{L}$  distinct interaction topologies. For the IBM simulations we choose a symmetrical network with four links, shown as M1. It was experimentally determined that the symmetry of this network promoted stability in the population dynamics (results not shown). For example the introduction of the link  $\hat{J}_{03}$  to M1 makes the network asymmetric, benefiting species 3 over species 1 (two feeding links versus one), and benefiting species 2 over species 0 (one predator versus two). The network M1 contains two plants (species 0 and 2), two herbivores (species 1 and 3) and a single predator (species 4).



Five species IBM dynamics with interaction topology M1 are shown in figure 5.38. Panels A and B show dynamics at low IR, which are more variable than the high IR dynamics shown in panels C and D. In what follows we focus on the high IR case based on previous observations that it is easier to fit to dynamics that are less complex<sup>8</sup>. This choice reflects an attempt to slightly simplify the challenge posed by the inference of interactions from the 5 species system. We again

<sup>8</sup>Comment on this in the three species case, regarding high versus low IR?

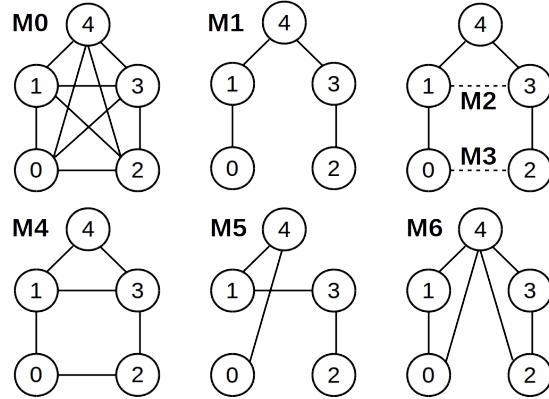


Figure 5.37: Seven candidate 5 species models. M0: fully connected. M1: the true topology of the IBM simulations to which the GLV is fitted. M2 and M3: One link added to M1, as indicated. M4-6: Different but plausible topologies. All models include intra-specific interactions (self-loops), although these are not drawn.

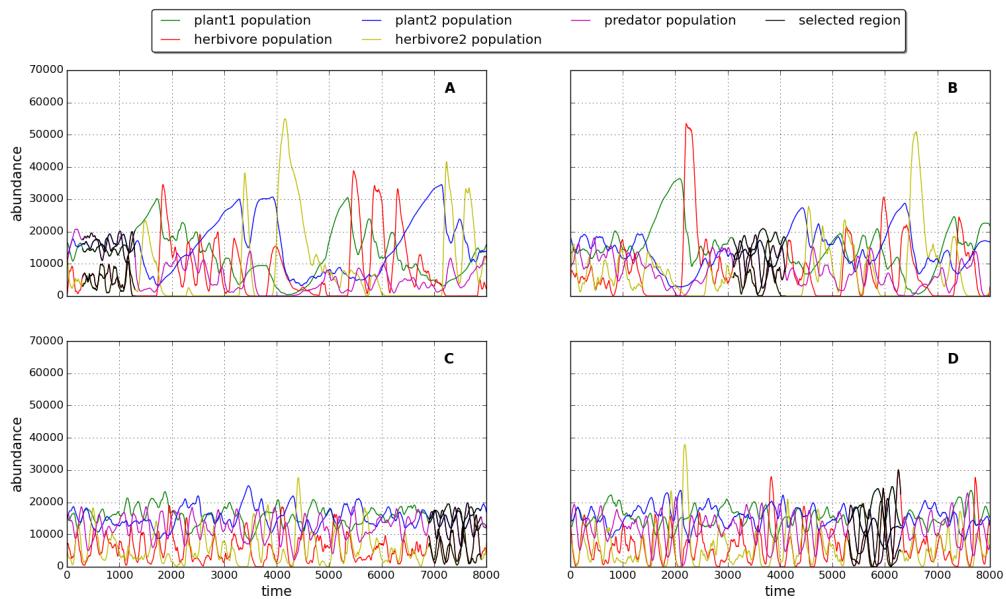


Figure 5.38: Similar to figure 5.30, but for five species simulations of the IBM with interaction topology M1 (shown in figure 5.37). Panels A and B: low IR. Panel C and D: high IR. The regions plotted in black represent selected *data streams* to which the GLV is fitted (selected using same criteria as in section 5.4.4).

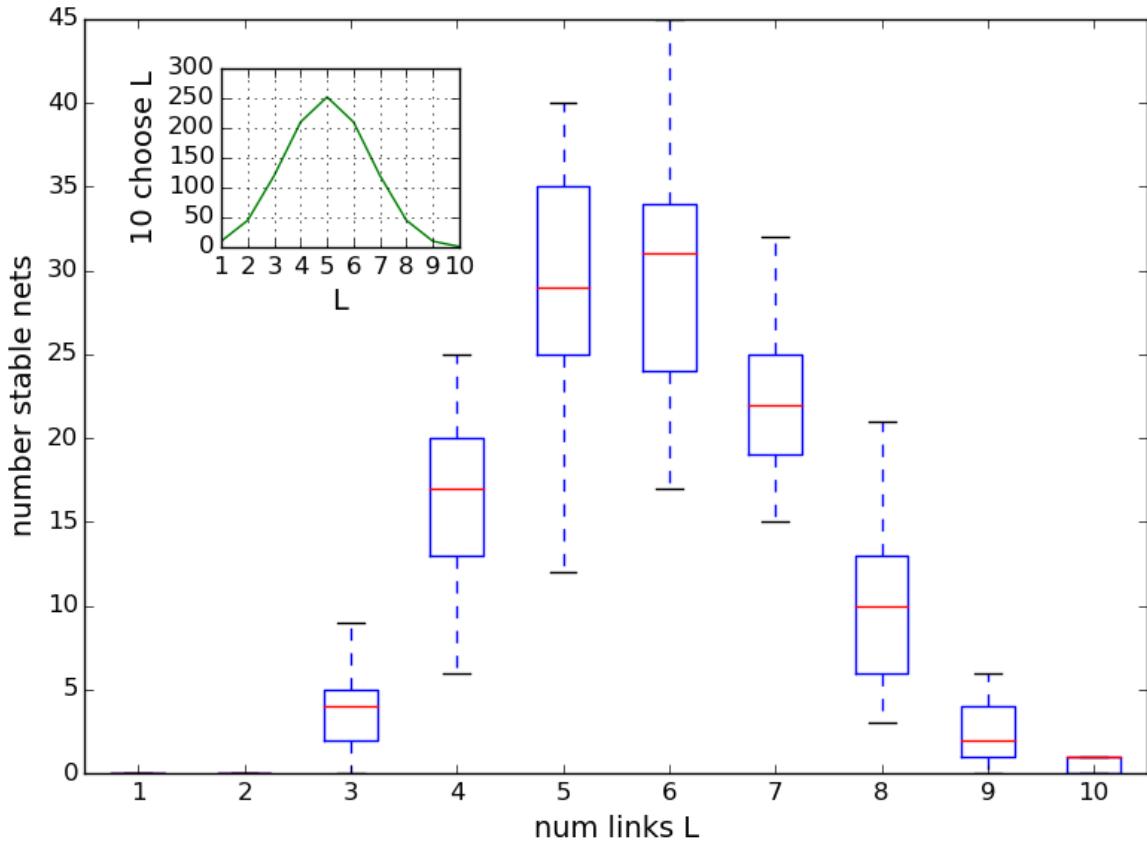


Figure 5.39: **The number of stable 5 species topologies** with  $L$  inter-specific interactions, when fitted to 25 replicate IBM simulations at **high IR**. The IBM simulations use topology M1 (figure 5.37), which has  $L=4$ . The inset shows the total number of possible networks with  $L$  links, all were tested. Stability is defined as the local stability of the equilibrium of the fitted model.

select the optimum region of 1000 time steps from the dynamics to use as the data stream. The regions are selected in the same way as in section 5.4.4, and are illustrated in figure 5.38. In this analysis we focus on the seven models depicted in figure 5.37. M0 is the full unconstrained GLV topology, while M1 is the true topology as discussed. The other models are selected as plausible alternatives to M1<sup>9</sup>.

Given the large number of possible topologies we first study the stability properties of all the potential competing models. Figure 5.39 summarises how many stable models exist at each value of  $L$ , for all 25 repeat simulations. Stability is again defined by the Jacobian of the fitted models. The number of stable models varies with  $L$  in a manner the approximately matches the total number of possible models (shown as an inset in the figure). However the most possible models exist at  $L = 5$ , whereas the greatest number of stable models exist at  $L = 6$ . The true model M1

<sup>9</sup>M2 and M3 by allowing possible competitions?

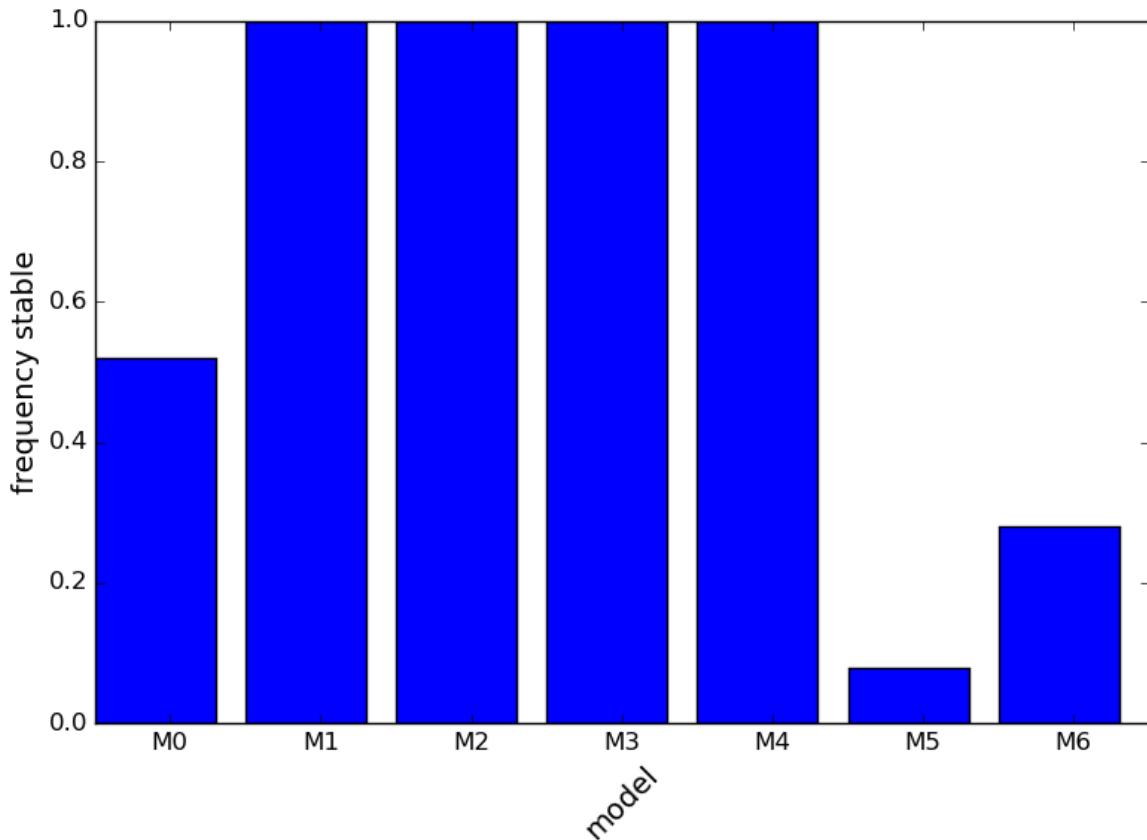


Figure 5.40: **The frequency with which each model (M0-6) is stable** when fitted to 5 species dynamics at **high IR** (25 replicate simulations). Stability is defined as the local stability of the equilibrium of the fitted model. The models M0-6 are shown in figure 5.37.

contains four links. At this value of  $L$  there are  $\binom{10}{4=210}$  possible models, of which only between 5 and 25 are dynamically stable. This raises the question as to whether M1 belongs to the stable set of models with  $L = 4$ . Figure 5.40 shows the frequency with which each model M0-6 is stable over the 25 replicates. We see that M1 is stable in all 25 cases, as are M2, M3 and M4. The latter models represent M1 with addition of either one or two links between species in the same trophic level (plant-plant, herbivore-herbivore, or both). We discuss the frequent stability of these three models further in section 5.5<sup>10</sup>. Model M0 is stable in around half of the simulations, while models M5 and M6 are rarely stable. Therefore it appears that focus on dynamics stability can indeed reduce the search space for identification of the correct topology, especially if the number of links in the true network is known.

Figure 5.41 summarise the relative errors in the demographic rate predictions for the seven models M0-6, and an additional two models. These models, M7 and M8, are selected on a case by

---

<sup>10</sup>Or in final section with plots to show competition?

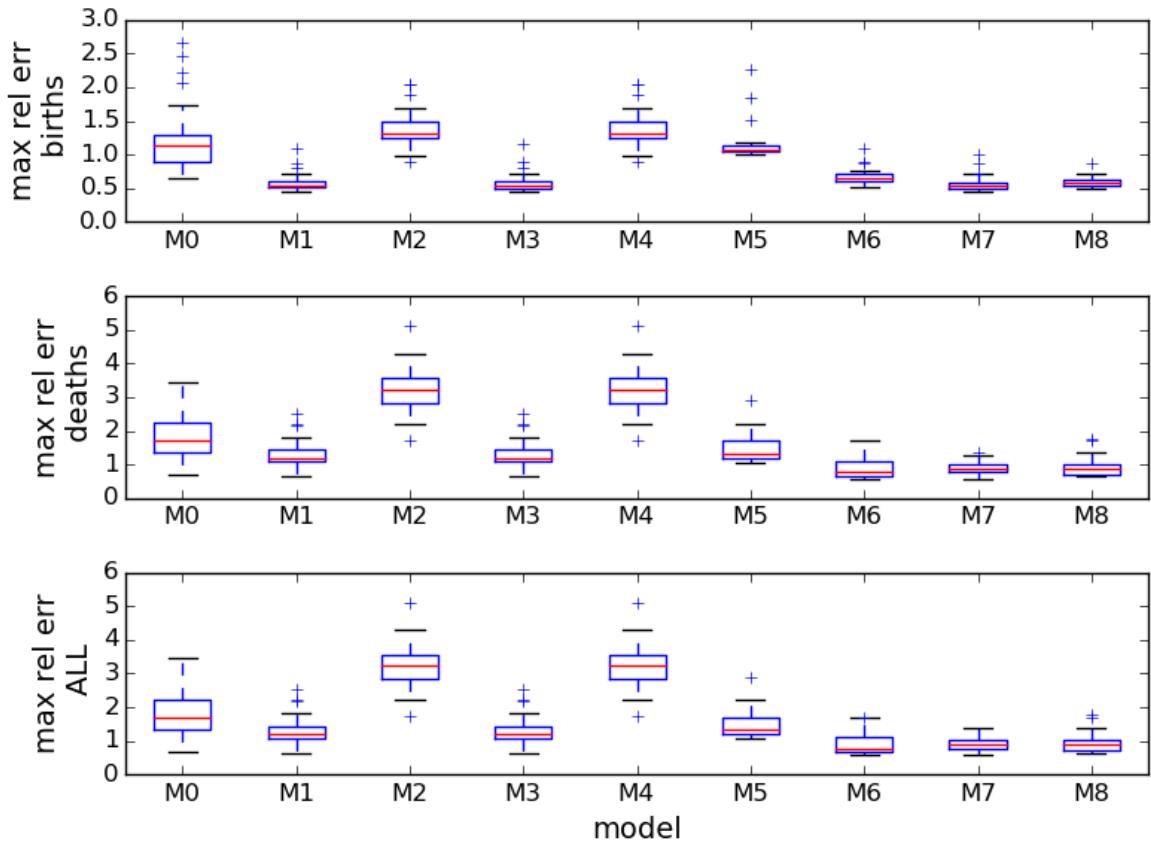


Figure 5.41: **Relative errors in the demographic rate predictions** of the models M0-6 (figure 5.37), plus two additional models: M7 and M8. M7 is selected as the model in the set of stable models with four links ( $L=4$ ) that has the lowest relative error. M8 is the same but for  $L=5$ . As such the topologies of M7 and M8 are not fixed. The errors shown are the maximum relative error in the predicted births (top row) and predicted deaths (middle row) for each simulations, and the maximum relative error in all rate predictions (bottom row). Relative error is defined in equation (5.41), and example rate predictions for three species are illustrated in figure 5.31.

case basis from the set of stable models with  $L=4$  and  $L=5$  respectively. The models are selected such that the maximum relative error in any rate prediction is minimised. Therefore the topology of these models can vary between simulations. From the figure we observe that, on aggregate, the predictions of M1, M3 and M5 are of comparable quality. The models M6-8 perform better, but only slightly. All models have a median maximum relative error (bottom panel) that is approximately equal to or greater than one. Therefore none of the models give accurate rate predictions for all species. We acknowledge that some information is lost by the use of the maximum relative error metric. For example a model may produce very accurate predictions in all but one of the rates. Additionally we note that the use of greater  $L$  values increases the number degrees of freedom, and therefore is likely to improve fit of the constrained GLV. Therefore further analysis should employ more sophisticated techniques for model comparison, such as the *Akaike information criterion* (AIC) [1]. In general we conclude that the true model (M1) performs favourably compared to competing alternative models in terms of stability and rate predictions. However no consistent criteria for the selection of the correct network topology have been identified.

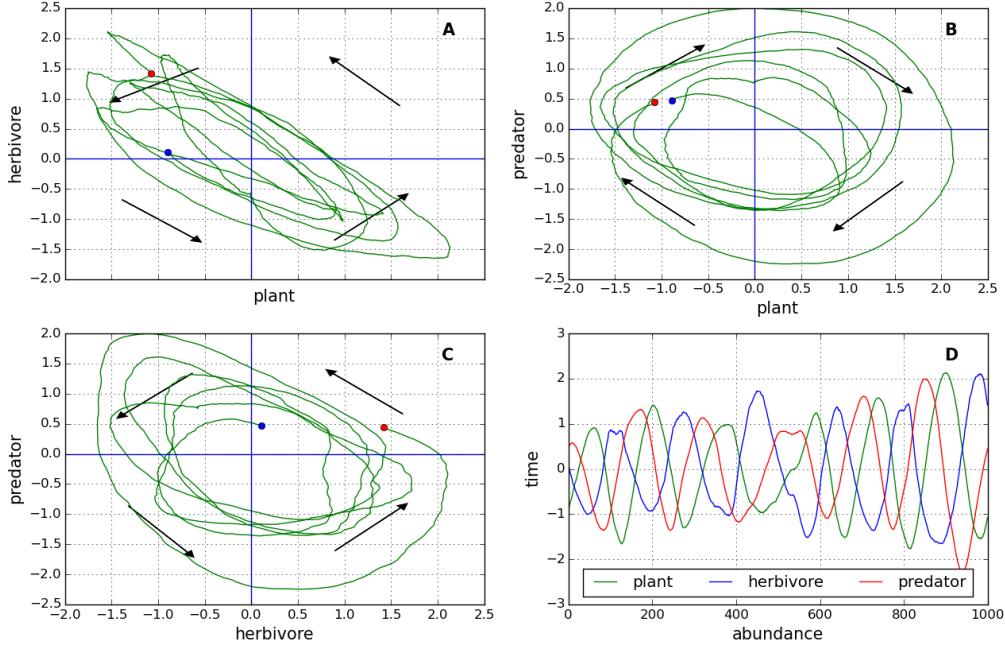
#### 5.4.6 Phase space analysis

One feature of the inference method, emerging from the three and five species analysis, is its tendency to detect spurious types of interaction between plant and predator species. In this three species case the estimated model parameter suggested that the predator was being eaten by the plant. In the five species case the same phenomenon was observed. When fitting M5 (from figure 5.37) the estimates  $\hat{J}_{04}$  and  $\hat{J}_{40}$  were invariably greater than and less than zero respectively (results not shown)<sup>11</sup>. In this section we propose an explanation for why the inference method produces this spurious result.

A robust feature of predator-prey dynamics, such as those modelled by the GLV, is that the oscillations of the predator population lag behind those of the prey population. The phase lag described results in trajectories with anti-clockwise rotation in the phase plane when the prey population is plotted on the x-axis, and the predator on the y-axis. This feature has been used to attempt to infer the presence of predator-prey interactions from population dynamics. For example Sandvik et al. [123] developed a method to quantify the direction and extent of this rotation, in order to detect signatures of predation from empirical population dynamics. However in certain situations the phase relationship may be reversed. For example Gilpin showed that certain sections of the famous Hudson Bay hare-lynx time series [] display clockwise rotation in the phase-plane. This led to the playful title of his publication: ‘*Do hares eat lynx?*’ [48]. Various models have managed to produce predator-prey dynamics that display such clockwise rotation, for example by including *time-delay* in the interaction []. However it could be that the clockwise rotation results from other factors. The hare-lynx system is embedded in a larger food web []. It may be that trophic interactions with other species are enough to disrupt the usually robust

---

<sup>11</sup>Show them!



**Figure 5.42: Phase plane projections** of 3 species food chain dynamics of the IBM model. Anti-clockwise rotation in the phase plane is characteristic of the species on the x-axis being the prey. Dynamics are standardised as described in the text. Panels A-C: projections onto phase plane defined by the three pairs of species. Arrows indicate average direction of trajectories in each quadrant. Blue and red circles show begging and end of trajectories respectively. Panel D: The standardised three species dynamics. This is the same dynamics as the black region in panel A of figure 5.30 (i.e. low IR).

anti-clockwise rotation. Alternatively the dynamics may be disrupted by environmental forcing, which Mutshida et al. [102] suggest have a stronger effect on population dynamics than trophic interactions in natural systems.

In figure 5.42 we employ a simplified version of Sandvik's method [123] to study the dynamics of the 3 species IBM in the phase plane. We use the portion of low IR dynamics from panel A of figure 5.30 that was used to fit the GLV model in that case (region plotted in black). The dynamics is standardised by subtracting the mean and dividing each species population time series by its standard deviation. As such the trajectories are centred on the origin when projected in each phase-plane. The average direction of the trajectories in each quadrant are then calculated. The figure reveals anti-clockwise rotation in the plant-herbivore (panel A) and herbivore-predator (panel C) phase plane. As discussed these rotations are the standard characteristic of the trophic interactions involved between these two pairs of species. However the plant-predator phase plane

(panel B) displays clockwise rotation, which is characteristic of the plant eating the predator. We suggest that this rotation is the cause of the spurious link detected by the GLV fit. It may be that the robust phase relationship between plant and predator is the result of some underlying feature of the IBM dynamics, which the GLV does not model (for example time delay). Alternatively it may simply be the result of the incremental phase lags produced by the two interactions depicted in panels A and C. In either case this result represents a simple proof that reliance on phase relationships for the detection of species interactions in multi-trophic communities is unwise. The implications of this result for our inference method are discussed further in the next section (5.5).

## 5.5 Discussion

In this section we summarise the main results of the chapter and discuss the limitations that have been discovered.

- Method works well for two species. Both ODE and IBM cases.
- Range-sampling method looks interesting, but is highly sensitive to noise. Can this be improved on? More sophisticated binning algorithm?
- In the case of three and five species the method produces reasonable predictions of demographic rates. The use of the maximum error may be a poor metric for comparison.
- We were not able to consistently detect the correct topology for the three and five species IBM dynamics. However the stability of the fitted model emerged as a significant feature.
- A more sophisticated model selection method (AIC) may tell us more about: A) the quality of the predict rates, and B) the correct topology (based on a range of features).
- In general the method struggles to fit the GLV to complex dynamics. And is hindered by noise.
- Are these the due to inadequacies of the GLV model, or in the model fitting method. Likely a combination. Analysis using an alternative method to fit the GLV model could reveal this.
- The sensitivity of the method to spurious phase relationships appears a major weakness. Possibly the fact the each species is fitted separately is the reason for this. Propose an alternative method: s-step ahead method [70] that may improve this because it involves simulating the fitted model.

## 5.6 Further work

In this section we consider further developments of the methodology, and the scope for application to empirical data.

In some ways the use of the method presented this chapter represents a naive treatment of the subject. We assumed no *a priori* knowledge about the *data generator* that produced the dynamics. In empirical applications the ecologist would have access to some information about the system. Therefore the results of the method could be improved by taking an informed approach to inferring species interactions. In the example of a marine plankton community, it may not be clear which species are interacting, but approximate species roles (e.g. autotroph versus heterotroph) and trophic positions may be known. Therefore some constraints can be placed on the GLV fit. Additionally the results of the inference method can be compared with previous knowledge, in order to determine plausibility. For example we have seen that there are often only a small number of stable competing topologies (for five species). Therefore a reasoned comparison of these topologies may reveal which one is most likely to be correct.

An additional result from this chapter is the relative success of demographic rate predictions by the fitted models. The predictions provide an alternative way to tests the methodology against empirical data. This would require a dataset with high resolution time series, and quantification of biomass/carbon flows between species. An candidate dataset for such a study comes from Lake Constance [].

Theoretically our investigation made certain modelling assumptions, which may need to be altered before empirical application. In particular *seasonality* was neglected. In nature the influence of environmental variables cannot be ignored, and indeed it has been suggested that they are more important in driving population dynamics than species interactions [102]. Therefore a natural extension of our methodology would be to consider seasonality. Also we have seen that the functional responses in the IBM are approximately linear. It would be informative to include prey-handling times into this model, and determine the effect that this has on our estimates.

Use other methods to constrain topology e.g. [55] can detect which interactions are direct

In the context of inferring gene regulatory networks, Marbach et al. [85] found that no single method (out of 30) performed optimally on all datasets tested. They concluded that the best practice was to use complementary methods for network detection to allow cross-validation. We propose that the same may be true for inferring species interactions. Therefore the most fruitful direction for future research may to studies that integrate and compare alternative methods.