***Supplementary Material***

**S1. Computational algorithm of ARZIMM**

Maximization of the penalized log-likelihood function corresponding to equation (2) with respect to is a computationally challenging task. This is mainly because both the integral with respect to the random effects and the zero-inflated structure do not have analytical solutions.

The Expectation-Maximization (EM) algorithm [1] has been traditionally used under a zero-inflated model in which the labels of zeros are treated as "missing data". Define the Bernoulli random variable if is drawn from zero state, if is drawn from distribution . Thus the “complete” data for the EM algorithm are . Here, play the role of missing data, denoted as . Based on , the complete data likelihood is given by

where .

Laplace approximations are widely used for maximum likelihood estimation of parameters involving the integral of random effects. Generally, the Laplace approximation is used to calculate integrals of the form.

where , can be simply defined as in our optimization problem, is a known function of variable . Let then the Laplace approximation is given by where . By implementing a Laplace approximation to (2) that

the estimator can be obtained via EM algorithm that calculate the expectation of and compute parameters iteratively as algorithm 1.

|  |
| --- |
| **Algorithm 1** computational algorithm for ARZIMM |
| Given :  **Repeat** for *h*=0,1,..  E-step (Outer): calculate the expectation of .  **Repeat** for *s*=0,1,...  Given , where :  M-step (Inner): solve the following minimization problem:  .  E-step (Inner): estimate , that for m-th element    **until** convergence. Calculate the expectation of .  M-step (Outer): solve optimization problem to obtain the estimate    **until** convergence. |

**S2. Competing methods**

**S2.1 Multivariate autoregression method (MAR)**

Multivariate autoregression (MAR(1)) models are applied to estimate the ecological interactions after a log transformation of ecological time series counts[2], formulated as

where is a vector of (log-transformed) population abundances at time , is a vector of constants, is a matrix whose elements give the effect of the abundance of species on the per capita population growth rate of species , and is a vector of process errors that has a multivariate normal distribution with mean vector and covariance matrix.

We implement the regularized maximum likelihood (ML) estimation procedure to select variables and estimate parameters simultaneously. Assuming that the true underlying effects are sparse, we advocate a Lasso-type approach, which adds an -penalty for the effects to the likelihood function. The tuning parameters are selected using Bayesian information criterion (BIC).

**S2.2 Poisson auto-regression method (Poisson)[2]**

In order to model the longitudinal sparse microbial counts from a study cohort, let denote the observed count of bacterial taxon for subject at time point . Consider a Poisson distribution of microbial count that , where represents the Poisson distribution with the parameter . Considering the longitudinal microbial counts, the Poisson autoregression is performed as:

,

Where and .

Similar regularized ML estimation procedure is implemented by adding an -penalty for the effects to the likelihood function. Bayesian information criterion (BIC) is applied for the tuning parameters selection.

**S2.3 The Microbial Dynamical Systems INference Engine (MDSINE)**

To infer dynamical system models of the microbiota from high-throughput time-series data.The Microbial Dynamical Systems INference Engine (MDSINE)[3] software package implements the extended generalized Lotka-Volterra (gLV) equations model for taxa measured in subjects, the rate of change of the concentration of taxon in subject , expressed as:

where represent unbounded growth rates, the represent pair-wise microbe-microbe interactions, and the represent effects of perturbations. The functions are binary valued, indicating if the given perturbation is present at time . represents an estimate of the concentration of taxon in subject at time point .

The MDSINE software is is publicly available [3]

**S3. Ergodicity and stability measures**

**S3.1 Ergodicity**

In this section, we establish the existence and uniqueness of the invariant distribution of the observation process based on ARIZMM, and shows the ergodicity and existence of some moments for its stationary distribution. The results are formed under the assumption that the time independent parameters and are known. To derive the general conditions, we introduce an unobserved Markov process based on a measurable function and a Markov kernel from to , which satisfies the following recursions:

(S1)

where denotes the Borel -algebra for . Specifically, is defined as satisfying (3) and is a zero-inflated Poisson distribution satisfying (1) where each dimension is independently distributed as a zero-inflated Poisson with parameters . We will give the assumptions so that the unobserved Markov process converge to the stationary distribution. Then, we will show the existence of a strict-sense stationary ergodic process on , the solution to the recursion (S1). Finally, we will derive the conditions meeting the assumptions for stationarity.

Based on the theory of Markov chains without irreducibility assumption to prove the existence of a stationary distribution, the existence of a stationary distribution has been investigated for the log-linear Poisson autoregression model[4]. We follow this approach and extend it to a zero-inflated multivariate version.

For a general Markov chain on state space with -algebra define for to be the -step transition probability starting from state . Define for all , as the law of that

(S2)

where . Consider the following assumptions:

(A1) The Markov kernel is a Feller semigroup. Moreover, there exist a compact set , ++ and a function + such that

(A2) The Markov kernel has a reachable point.

(A3) There exist a kernel on , a measurable function , and a measurable function such that for all ,

where denotes the euclidean distance. There exists such that for all

Moreover, for all , there exists such that

where is the open ball of radius with respect to and centered at .

**Theorem 1.** *Assume that all eigenvalues of matrix lie inside the unit circle. Then, the assumptions (A1)—(A3) hold.*

The proof can be easily adapted from [3].

**Theorem 2.** *Assume that (A1)—(A3) hold. Then, the Markov kernel admits a unique invariant probability measure.*

**Lemma 1.** *Assume that the Markov kernel admits an invariant kernel and that there exist a measurable function + and real numbers such that . Then,*

**Proposition 1.** *Assume that the Markov kernel admits a unique invariant probability measure. Then, there exists a strict-sense stationary ergodic process on ,, the solution to the recursion of the model.*

*Proof.* The property of asymptotically strong Feller Markov chains has been investigated in Hairer et al. [5]. The following theorem is taken from Theorem 3.16 in Hairer et al. [5].

**Definition 1.** *A Markov transition semigroup on a Polish space is called asymptotically strong Feller at if there exists a totally separating system of pseudo-metrics for and a sequence such that*

(S3)

**Theorem 3.** *Assume that is Markov semigroup satisfying (S3) and admits reachable points . Then there exists at most one invariant probability measure for .*

**S3.2 Stability measures**

**Return rate**

The first stability measure is the return rate. For a system, where a stationary distribution exists, the more rapid the convergence, the more stable the system. Following Ives et. al [2], we measure the rate of return by the rate at which the transition distribution converges to the stationary distribution from an initial condition . Previously, we established the existence of the stationary distribution for the observation process based on an unobserved Markov process with a stationary distribution. Therefore, we consider the return rate of the unobserved Markov process . Similar to the MAR(1) process, the return rate of the mean vector of is governed by , and limits the return time of the mean to .

**Reactivity**

The second stability measure is the reactivity, which assesses how strongly population-level microbiome abundances are pulled towards the mean of the stationary distribution. To quantify the reactivity, Ives et. al [2] considered the expectation over the stationary distribution of the change in distance from the mean of the stationary distribution between successive time steps. Similar to return rate, we consider the reactivity of the unobserved Markov process . Let be the difference between and the mean of the stationary distribution. Therefore, the reactivity is defined as

,

where denotes the squared Euclidean distance. To obtain the reactivity of the system averaged over a long time period, we can let the distribution of be the stationary distribution , and thus, the reactivity can be shown as

where is the dominant eigenvalue of the matrix . Thus, indicates the "worst-case" reactivity of a system.

# S4. Simulation setting

# S4.1. Simulation setting for the zero part

We generated the longitudinal absolute counts from zero-inflated Poisson distribution with parameters and for each taxon. Since our focus is on the estimation of the interaction matrix which depends on the non-zero part, we adopted a simple simulation design for the zero inflation proportions We ignored the individual variations in by dropping the random effect term in equation (1). In the simulation design, in order to mimic the observed sparsity in real data, we set the zero inflation proportions for 20 taxa as =. In order to test our algorithm on estimating the parameters in zero inflation part along with the non-inflation part, based on the model , we constructed six binary covariates with , and when , and when. The elements of were set to be sparse with informative coefficients to be 5.

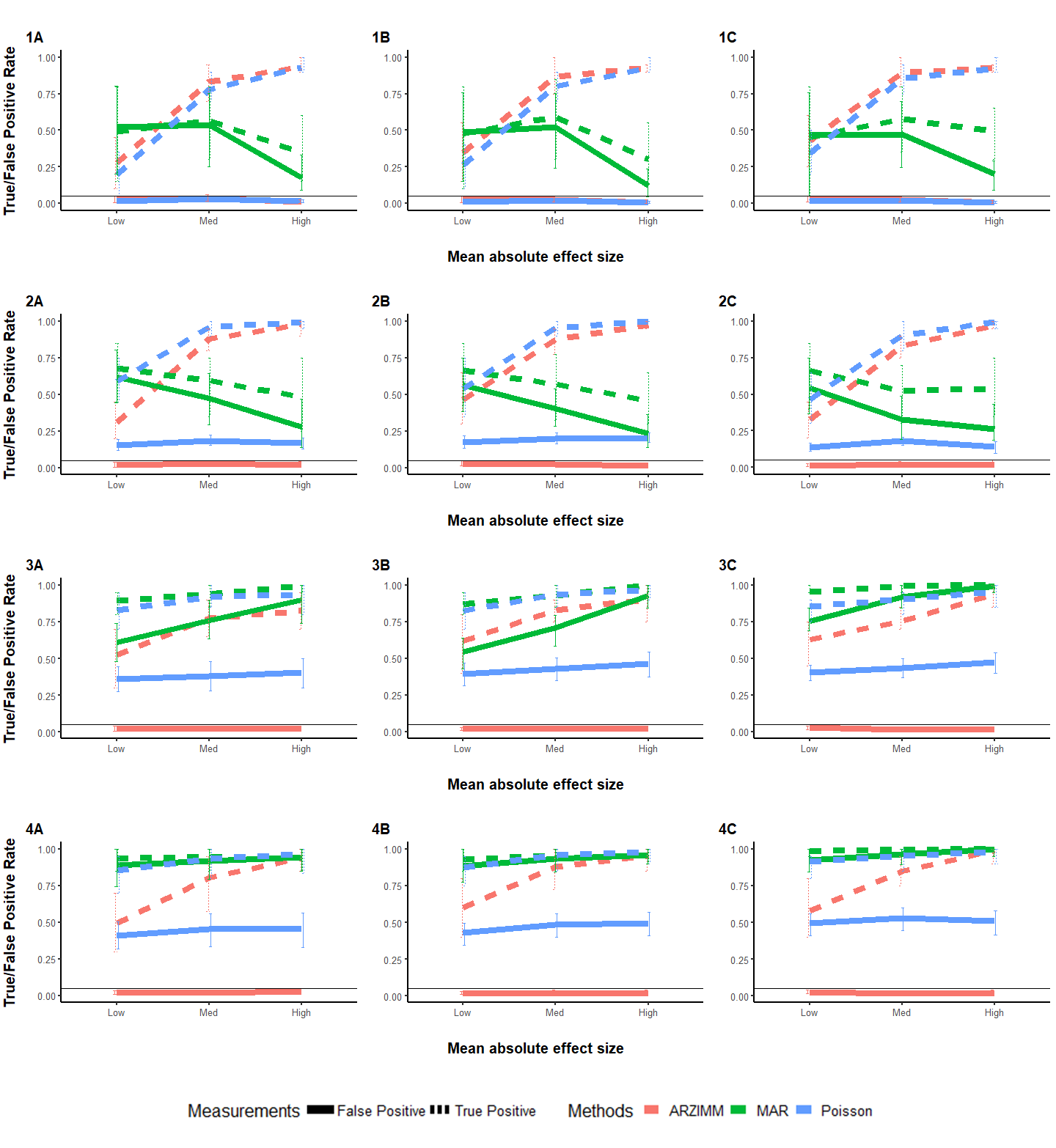
# S4.2. Simulation setting for the mis-specified model

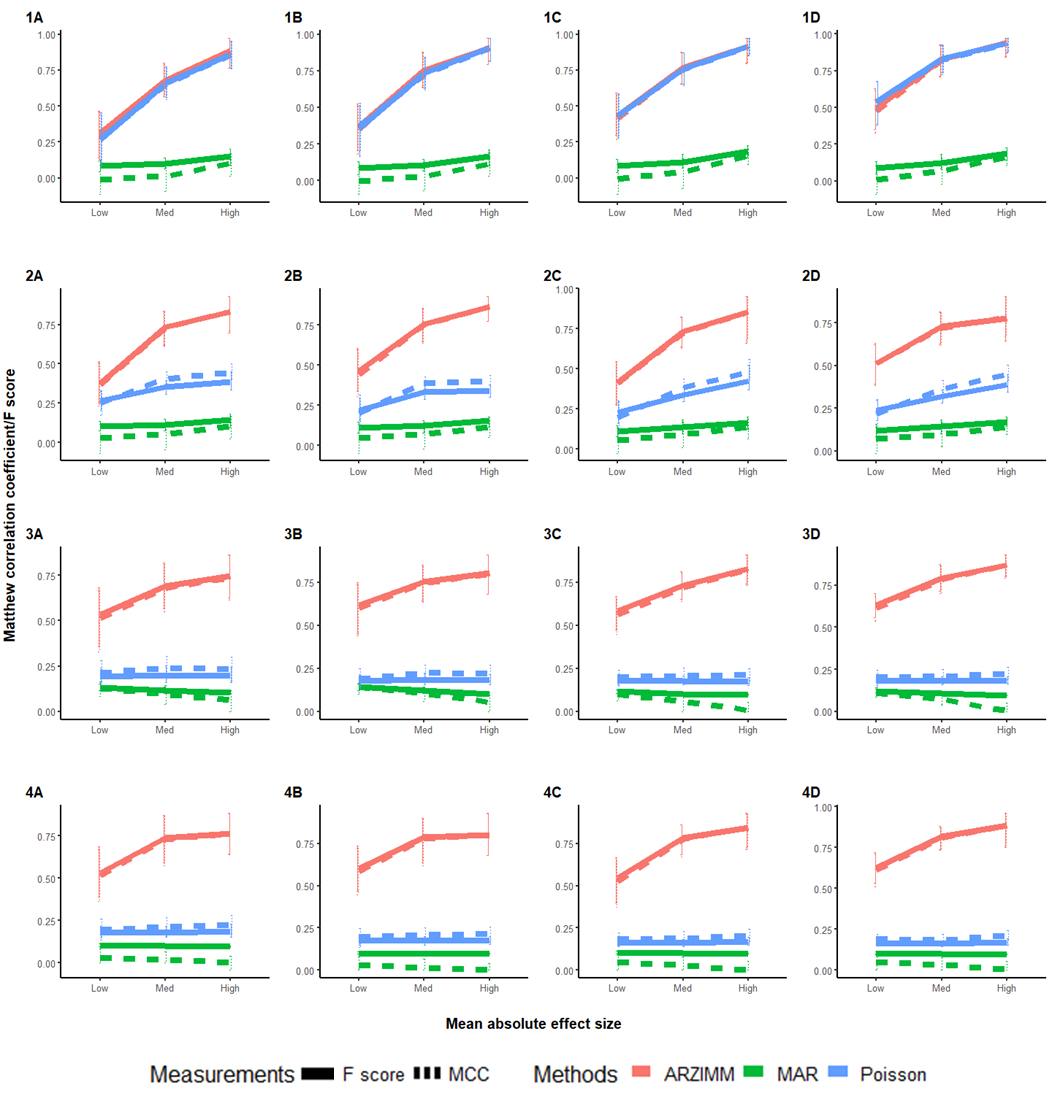
In this section, we conducted simulation experiments to verify the robustness of ARZIMM under model misspecification. Based on Scenario 4, where both zero-inflated structure and heterogeneity are modeled, we added a random term to the zero inflation proportions to allow the zero inflation proportions , where and ’s -th element (1) ; or (2) , to vary along the time. Here, , the strength of variation, was set based on the coefficient of variation estimates using MIME data as we reported in the Discussion. In this mis-specified simulation, we evaluated the variable selection performance for ARZIMM, Poisson, and MAR in terms of true positive rate (TPR) and false positive rate (FPR) and present the results in Supplementary Figure S3. The results show that the moderate temporal variation in probability of zero count does not affect ARIZMM’s performance much in capturing the informative interactions by estimating B when the absolute effect strengths of interaction matrices is high(FDR<0.05) or medium (FDR <0.15).

# 5. Supplementary Tables and Figures

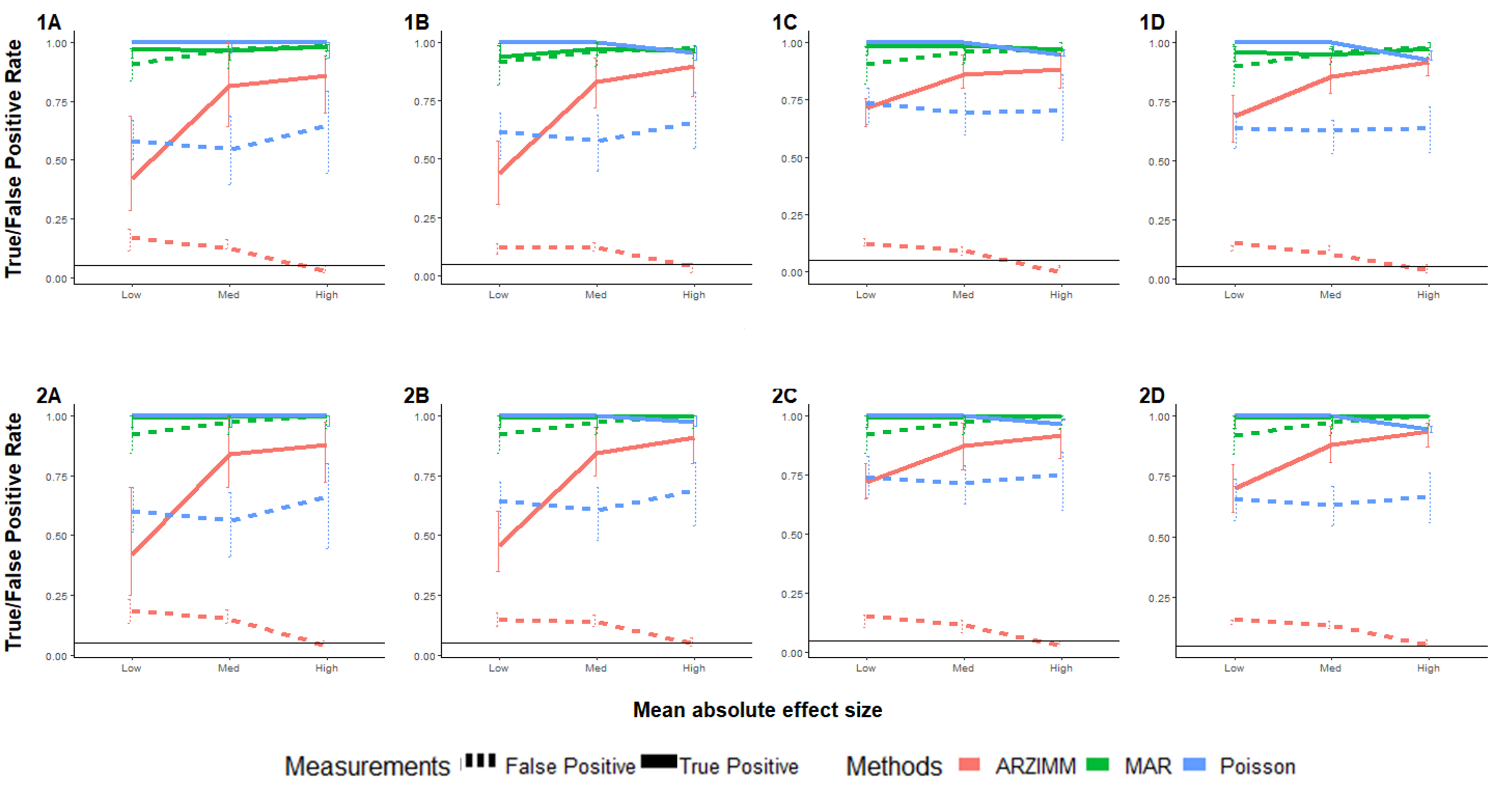
**Supplementary Table 1.** Simulation results for all settings under four scenarios. Poisson refers to the penalized Poisson auto-regression model and MAR refers to penalized log-normal multivariate auto-regression model. Reported value are medians (IQRs) of mean normalized squared error score (MNSES) calculated over 500 simulations for each setting.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Methods | | ARZIMM | | | Poisson | | | MAR | | |
| Effect size | | High | Median | Low | High | Median | Low | High | Median | Low |
| Scenario 1 | | | | | | | | | | |
| *n* | *T* | Median (IQR) | | | | | | | | |
| 20 | 20 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 76 | 74 | 74 |
|  |  | (0.98-0.99) | (0.98-1.00) | (0.99-1.00) | (0.99-1.01) | (0.99-1.00) | (0.99-1.01) | (53-120) | (53-122) | (53-131) |
| 50 | 10 | 0.98 | 0.98 | 0.98 | 1.01 | 1.00 | 1.00 | 78 | 77 | 84 |
|  |  | (0.98-0.99) | (0.98-0.99) | (0.98-0.99) | (1.00-1.01) | (0.99-1.00) | (0.99-1.00) | (55-133) | (55-134) | (59-139) |
| Scenario 2 | | | | | | | | | | |
| *n* | *T* | Median (IQR) | | | | | | | | |
| 20 | 10 | 0.97 | 0.98 | 0.98 | 1.51 | 1.33 | 1.23 | 70 | 56 | 57 |
|  |  | (0.96-1.48) | (0.96-1.09) | (0.97-0.99) | (1.50-1.53) | (1.32-1.35) | (1.22-1.24) | (49-109) | (40-86) | (38-94) |
| 20 | 20 | 0.99 | 1.06 | 1.00 | 1.49 | 1.34 | 1.25 | 91 | 80 | 79 |
|  |  | (0.98-1.36) | (0.98-1.08) | (0.99-1.01) | (1.48-1.50) | (1.33-1.35) | (1.23-1.26) | (66-156) | (57-130) | (57-135) |
| 50 | 10 | 0.96 | 1.09 | 1.17 | 1.57 | 1.36 | 1.26 | 89 | 78 | 80 |
|  |  | (0.96-0.97) | (1.05-1.15) | (1.14-1.21) | (1.56-1.62) | (1.36-1.37) | (1.26-1.27) | (63-148) | (54-124) | (55-127) |
| 50 | 20 | 0.98 | 1.13 | 1.14 | 1.54 | 1.36 | 1.26 | 117 | 116 | 114 |
|  |  | (0.98-0.98) | (1.07-1.17) | (1.11-1.19) | (1.53-1.56) | (1.36-1.37) | (1.26-1.26) | (82-191) | (79-209) | (79-179) |
| Scenario 3 | | | | | | | | | | |
| *n* | *T* | Median (IQR) | | | | | | | | |
| 20 | 10 | 0.92 | 0.87 | 0.88 | 27.6 | 19.4 | 19.6 | 35012 | 13069 | 9128 |
|  |  | (0.90-0.96) | (0.85-0.91) | (0.85-0.94) | (22.0-32.8) | (15.4-23.3) | (15.5-23.3) | (10154-101650) | (6552-24509) | (3320-16752) |
| 20 | 20 | 0.94 | 0.92 | 0.92 | 41.3 | 30.0 | 29.4 | 144005 | 32263 | 17706 |
|  |  | (0.92-1.30) | (0.90-0.94) | (0.90-0.93) | (35.6-45.4) | (26.5-32.5) | (26.4-31.6) | (64379-313924) | (20626-58587) | (10881-27450) |
| 50 | 10 | 0.95 | 0.95 | 0.95 | 38.5 | 28.0 | 26.8 | 131596 | 44419 | 24869 |
|  |  | (0.94-0.97) | (0.94-0.96) | (0.94-0.96) | (33.4-41.3) | (26.3-29.4) | (25.1-28.1) | (63797-242634) | (27851-83709) | (16184-44889) |
| 50 | 20 | 0.98 | 0.98 | 0.98 | 43.6 | 30.0 | 28.5 | 252042 | 63645 | 39205 |
|  |  | (0.97-0.98) | (0.97-0.98) | (0.97-0.98) | (42.0-44.8) | (29.3-30.8) | (27.6-29.1) | (164342-395273) | (46459-107196) | (27347-70133) |
| Scenario 4 | | | | | | | | | | |
| *n* | *T* | Median (IQR) | | | | | | | | |
| 20 | 20 | 0.95 | 0.92 | 0.84 | 43.9 | 28.5 | 28.7 | 150843 | 85354 | 43717 |
|  |  | (0.91-2.32) | (0.89-0.96) | (0.83-0.86) | (32.5-51.4) | (25.6-31.0) | (25.0-30.9) | (55236-446842) | (45584-169844) | (25602-83638) |
| 50 | 10 | 1.32 | 0.87 | 0.84 | 36.5 | 28.3 | 28.0 | 55616 | 67888 | 48800 |
|  |  | (1.24-1.41) | (0.86-0.89) | (0.83-0.86) | (31.8-42.8) | (26.3-30.0) | (26.2-29.5) | (30732-105802) | (43224-120112) | (33059-89420) |

**Supplementary Figure 1.** Simulation results of four scenarios for (A) 20 subjects with 10 time points; (B) 20 subjects with 10 time points; (C) 50 subjects with 10 time points. Poisson refers to the penalized Poisson auto-regression model and MAR refers to penalized log-normal multivariate auto-regression model. False Positive Rate and True Positive Rate are reported with mean and 95% interval over 500 simulations for four scenarios to have: (1) no zero-inflated structure and no heterogeneity; (2) heterogeneity but no zero-inflated structure; (3) zero-inflated structure but no heterogeneity; (4) both zero-inflated structure and heterogeneity.

****

**Supplementary Figure 2.** Matthew correlation coefficient and F-score over 500 simulations for four scenarios for (A) 20 subjects with 10 time points; (B) 20 subjects with 10 time points; (C) 50 subjects with 10 time points; (D) 50 subjects with 20 time points. Poisson refers to the penalized Poisson auto-regression model and MAR refers to penalized log-normal multivariate auto-regression model. False Positive Rate and True Positive Rate are reported with mean and 95% interval over 500 simulations for four scenarios to have: (1) no zero-inflated structure and no heterogeneity; (2) heterogeneity but no zero-inflated structure; (3) zero-inflated structure but no heterogeneity; (4) both zero-inflated structure and heterogeneity.



**Supplementary Figure 3.** Simulation results of mis-specified model for (A) 20 subjects with 10 time points; (B) 20 subjects with 20 time points; (C) 50 subjects with 10 time points; and (D) 50 subjects with 20 time points. Instead of true zero inflation proportions , the mis-specified zero inflation proportions are used to simulate absolute abundances. Poisson refers to the penalized Poisson auto-regression model and MAR refers to penalized log-normal multivariate auto-regression model. False Positive Rate and True Positive Rate are reported with mean and 95% interval over 500 simulations for 2 scenarios to have and ’s -th element: (1) and (2) .

Reference:

1. Dempster AP, Laird NM, Rubin DBJJotRSSSB: **Maximum likelihood from incomplete data via the EM algorithm**. 1977, **39**(1):1-22.

2. Ives AR, Dennis B, Cottingham K, Carpenter SJEm: **Estimating community stability and ecological interactions from time‐series data**. 2003, **73**(2):301-330.

3. Bucci V, Tzen B, Li N, Simmons M, Tanoue T, Bogart E, Deng L, Yeliseyev V, Delaney ML, Liu QJGb: **MDSINE: Microbial Dynamical Systems INference Engine for microbiome time-series analyses**. 2016, **17**(1):1-17.

4. Douc R, Doukhan P, Moulines EJSP, Applications t: **Ergodicity of observation-driven time series models and consistency of the maximum likelihood estimator**. 2013, **123**(7):2620-2647.

5. Hairer M, Mattingly JCJAoM: **Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing**. 2006:993-1032.