IDM

Malaria Model Validation Results

Validation Report

06-06-2022
Eradication version: 2.20.5097.0
branch: Malaria-Ongoing(6ee63023a)
emodpy_malaria version: 1.1.19
Suite ID: 3be02fa2-bee5-ec11-a9f9-b88303911bc1

Table of contents

1. Introduction

1.1 Background

The goal of this report is to help users quickly identify whether updated versions of the malaria model are still well-calibrated to capture a range of relevant real-world malaria observations.

The figures and tables compare simulation output generated with a particular version of the Eradication.exe and of emodpy-malaria with 1) the simulation results generated by earlier versions of Eradication.exe and emodpy-malaria (the versions used to calibrate the model) and 2) reference datasets from real-world observations.

This report was generated by running the malaria model validation workflow available at https://github.com/InstituteforDiseaseModeling/malaria-model_validation. Additional information on the reference datasets and on the simulation assumptions are available from the repo in "Notes on reference datasets and simulation assumptions.docx," and instructions on how to re-run the validation comparisons are in the README file.

2. Results summary

2.1 Validation test performance

The table below shows the number of sites examined in each validation relationship and the fraction of sites that passed each of several tests for that validation relationship.

PLACEHOLDER TABLE: These are made up numbers!

## 2.2 Performance compared to model version from calibration

The number of sites where the current simulation match to the reference data was found to be substantially WORSE than the version used in calibration are shown in the table below.

PLACEHOLDER TABLE: These are made up numbers!

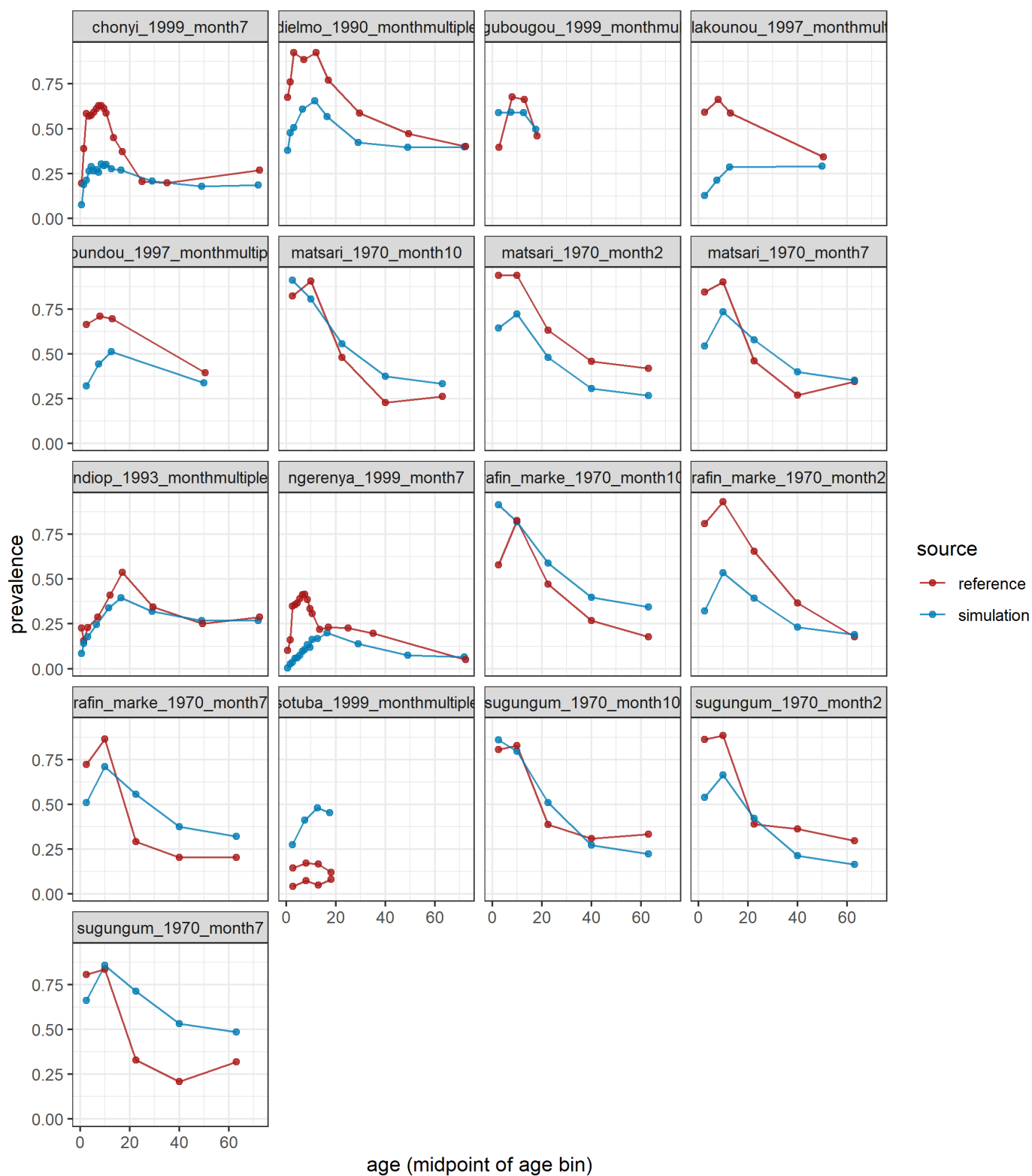# 3. Visual comparison of reference data and matched simulations

3.1 Incidence by age

The plots below compare the age-incidence relationships from reference datasets and matched simulations.
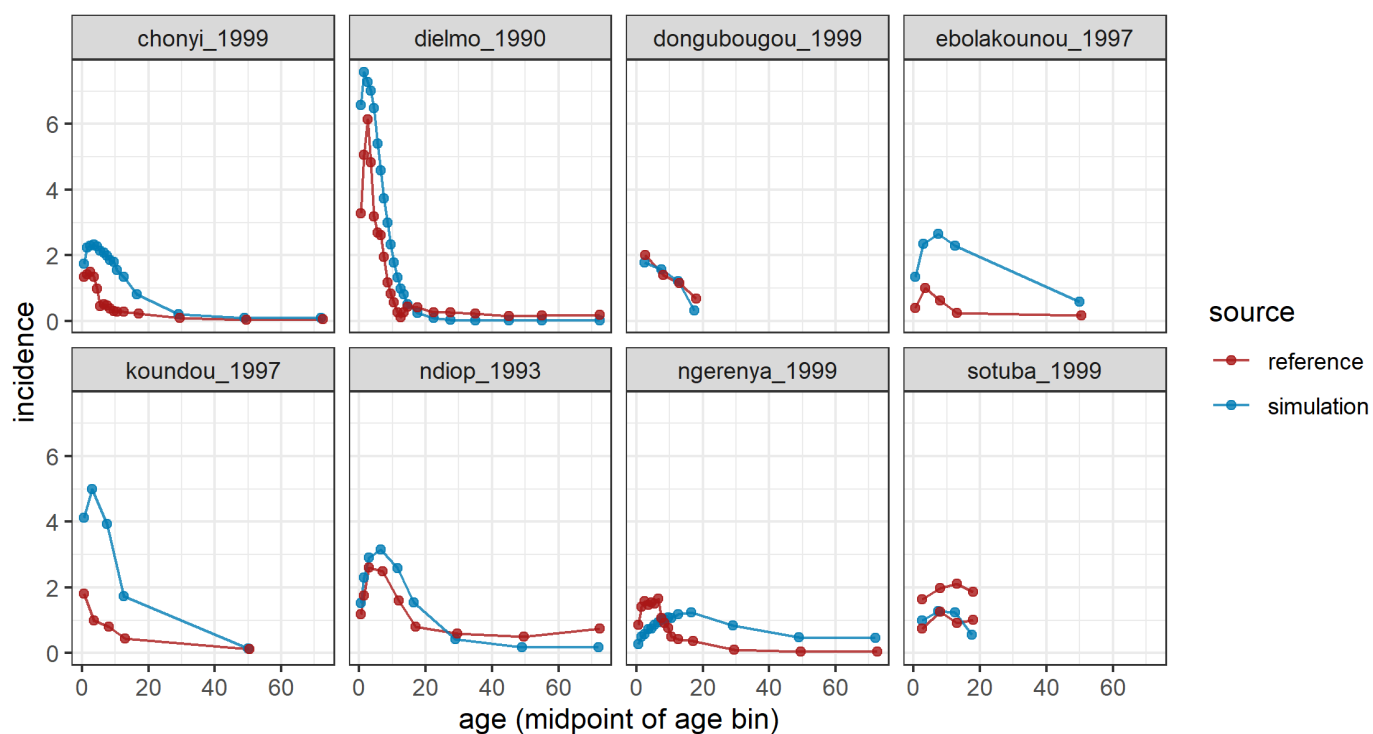
3.2 Prevalence by age

The plots below compare the age-prevalence relationships from reference datasets and matched simulations.

3.3 Infectiousness to vectors

Each of the below plot panels corresponds to a site. Within a plot panel, each row corresponds to an age group and each column corresponds to the month when sampling occurred.
The x-axis shows the gametocyte density in an infection. The y-axis shows how infectious an individual is to mosquitoes. The dot size shows how often a person of a given age and gametocyte density falls into each of the infectiousness bins (each column's dot sizes sum to one).
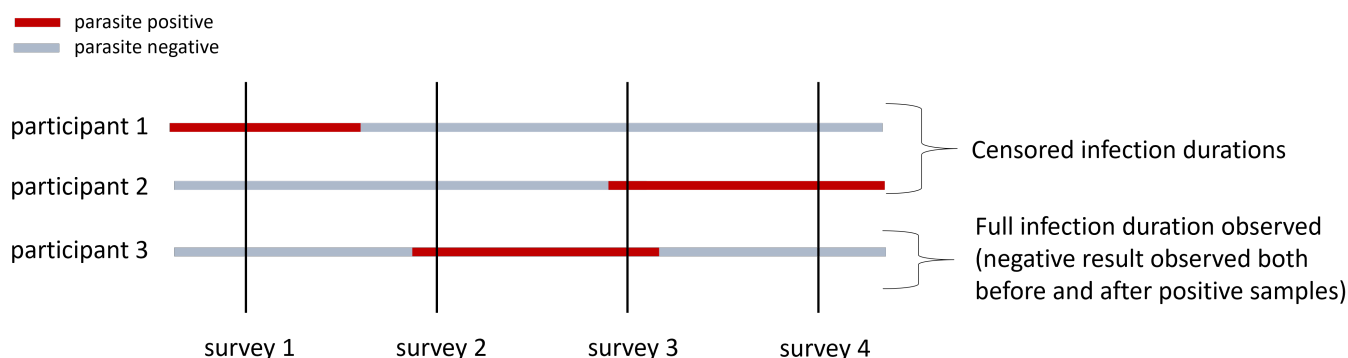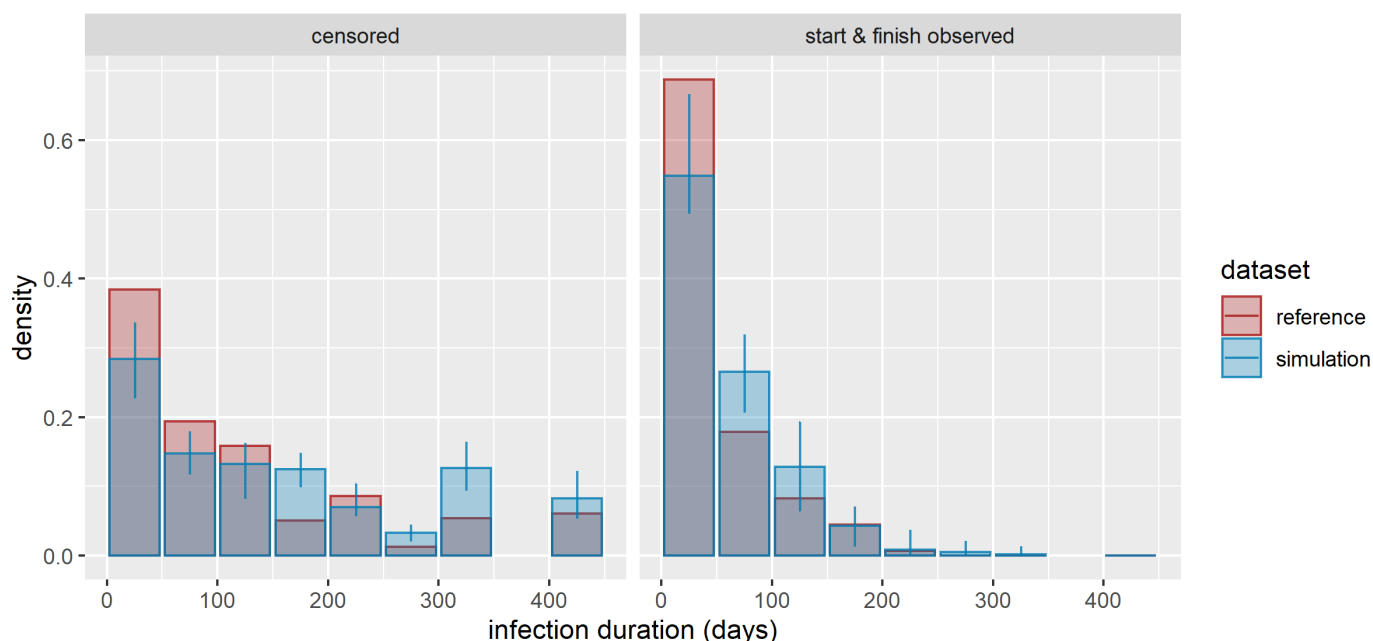In the reference datasets, the sample size is sometimes quite small.

dapelogo_2007

laye_2007
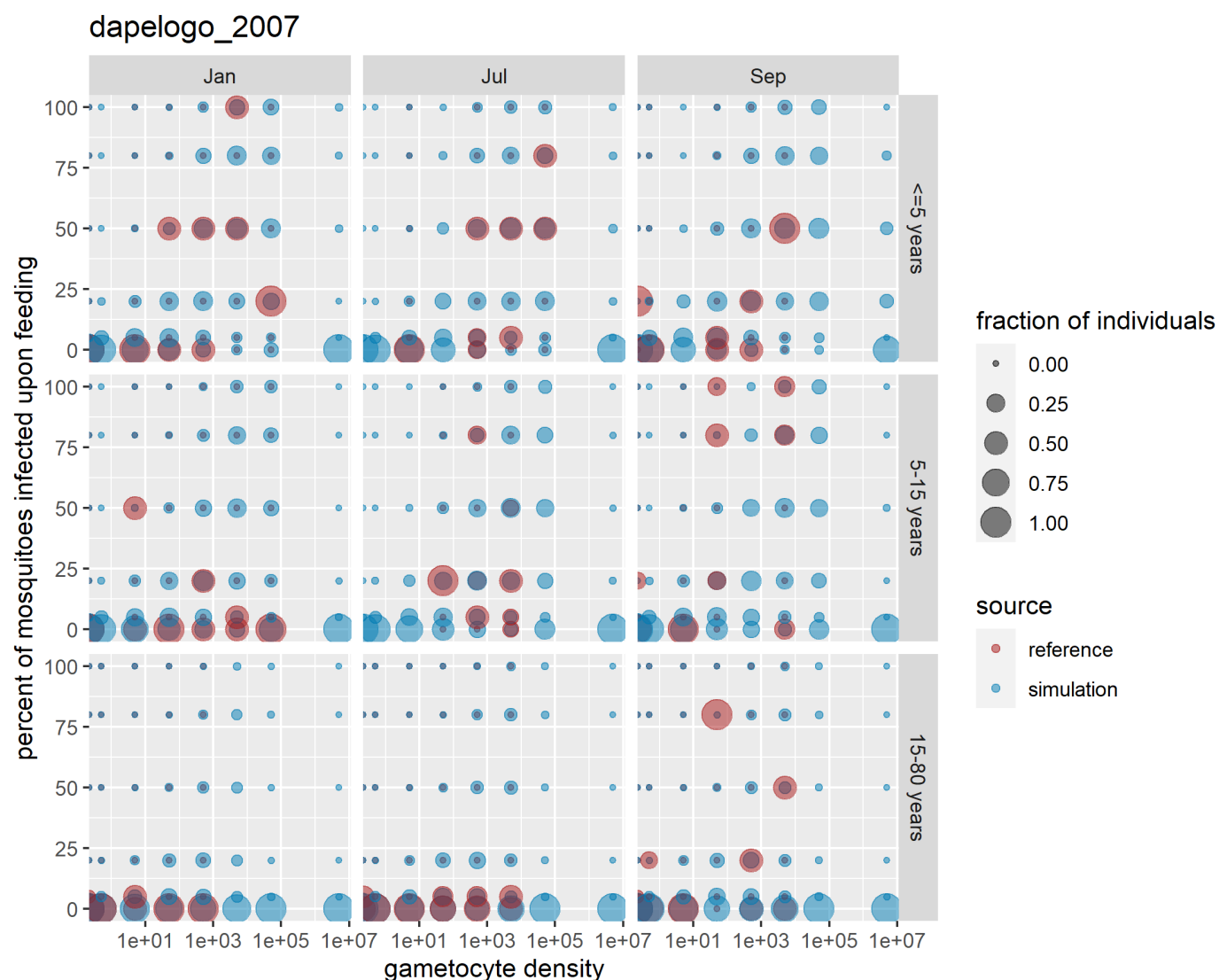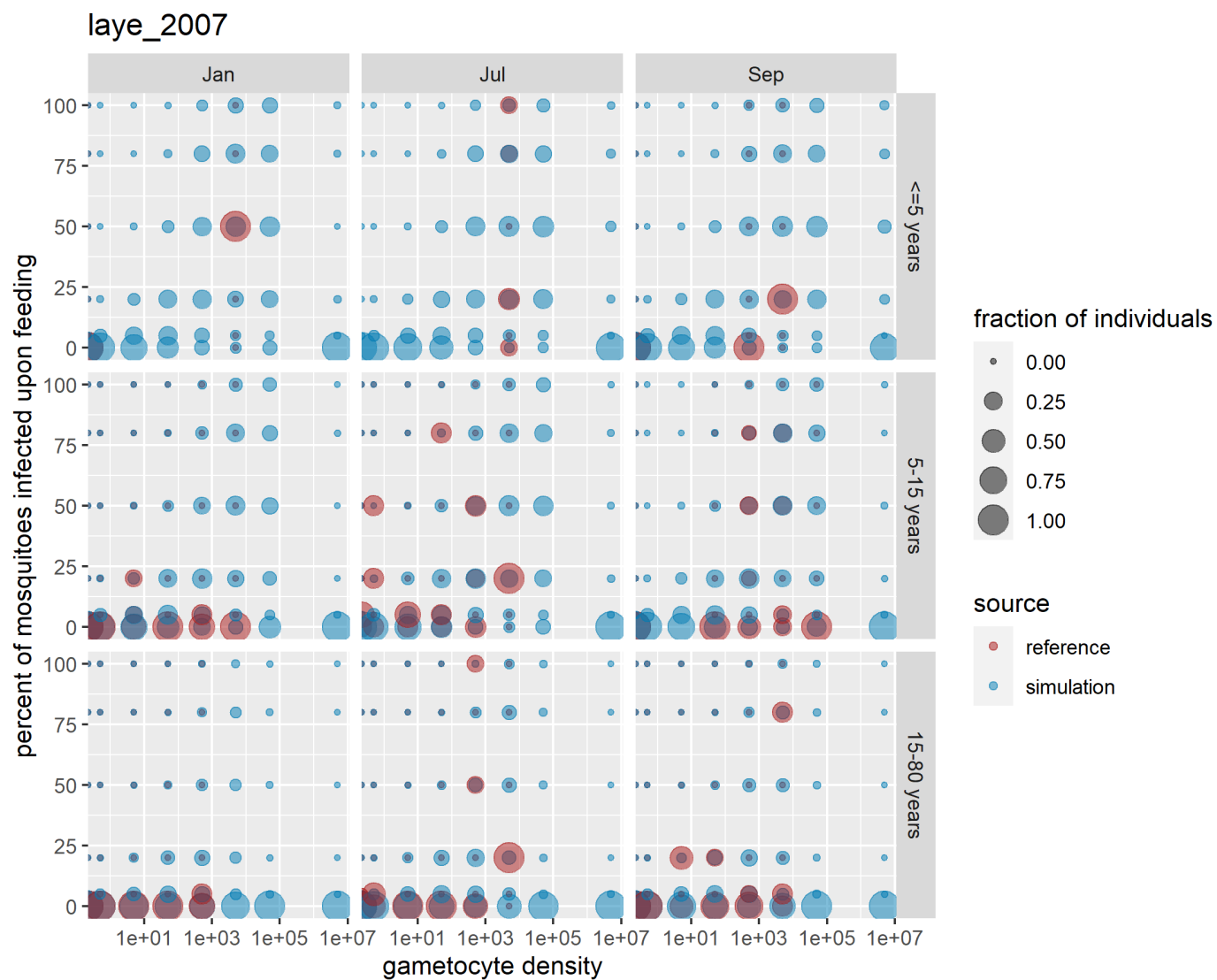
3.4 Duration of infection - all ages

The plots below compare the duration over which individuals had positive tests in the reference dataset and matched simulations. The sampling design from the reference data was matched in the simulations.Observed infections are divided into two groups. "Censored" infections refer to infections where the individual was positive at the first or final survey of the study (so the infection may have extended beyond the period observed). "Start & finish observed" infections refer to infections were the individual was observed to have a negative test at the start and end of the infection. The two types of infection duration records are illustrated in the figure below.
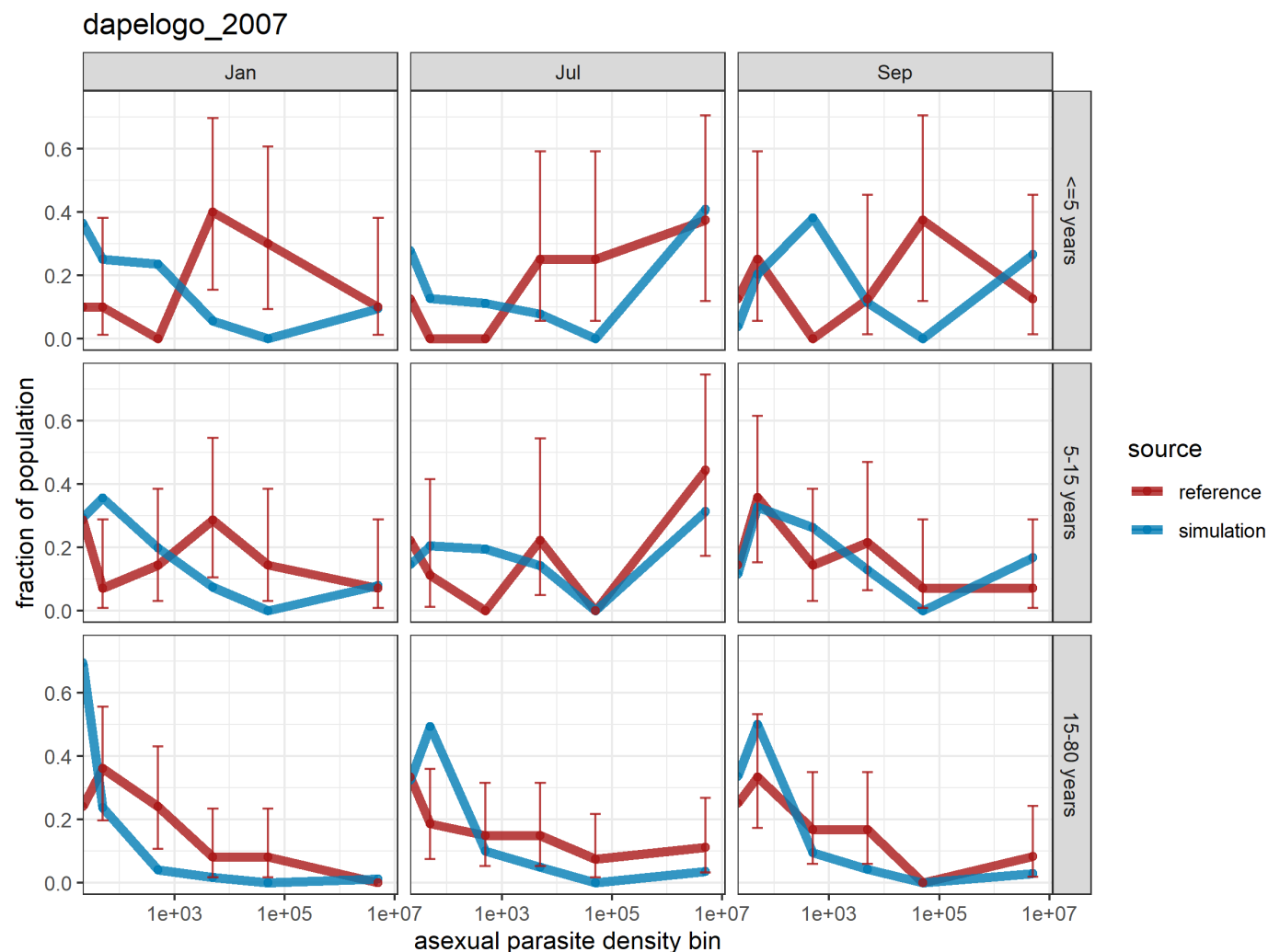
3.5 Duration of infection - by age

The plots below compare the duration over which individuals had positive tests in the reference dataset and matched simulations. The sampling design from the reference data was matched in the simulations.Observed infections are divided into two groups. "Censored" infections refer to infections where the individual was positive at the first or final survey of the study (so the infection may have extended beyond the period observed). "Start & finish observed" infections refer to infections were the individual was observed to have a negative test at the start and end of the infection. The two types of infection duration records are illustrated in the figure below.

In the plot panel below, columns correspond to the age group (in years) and rows correspond to whether or not the start and end of the infection was observed.
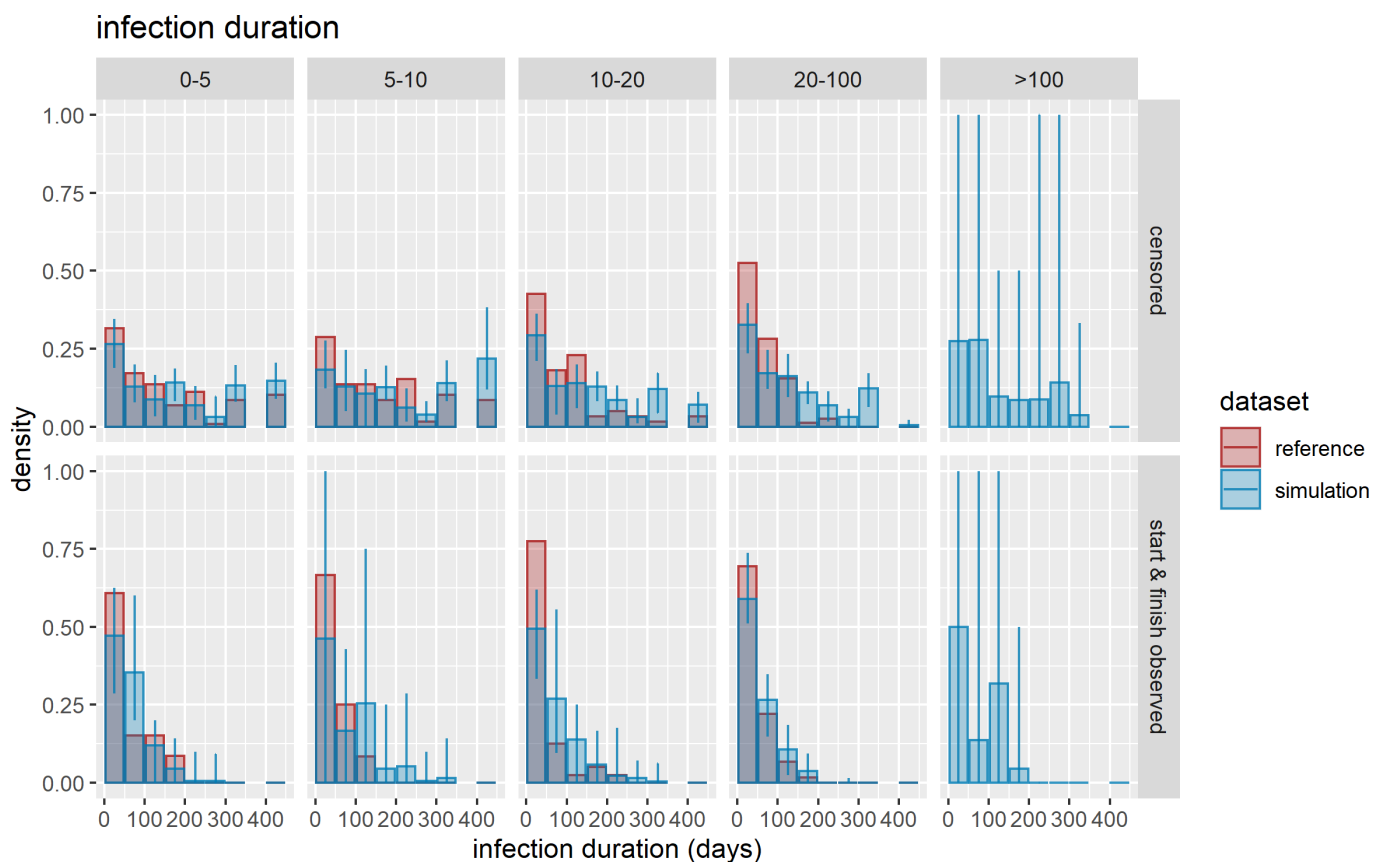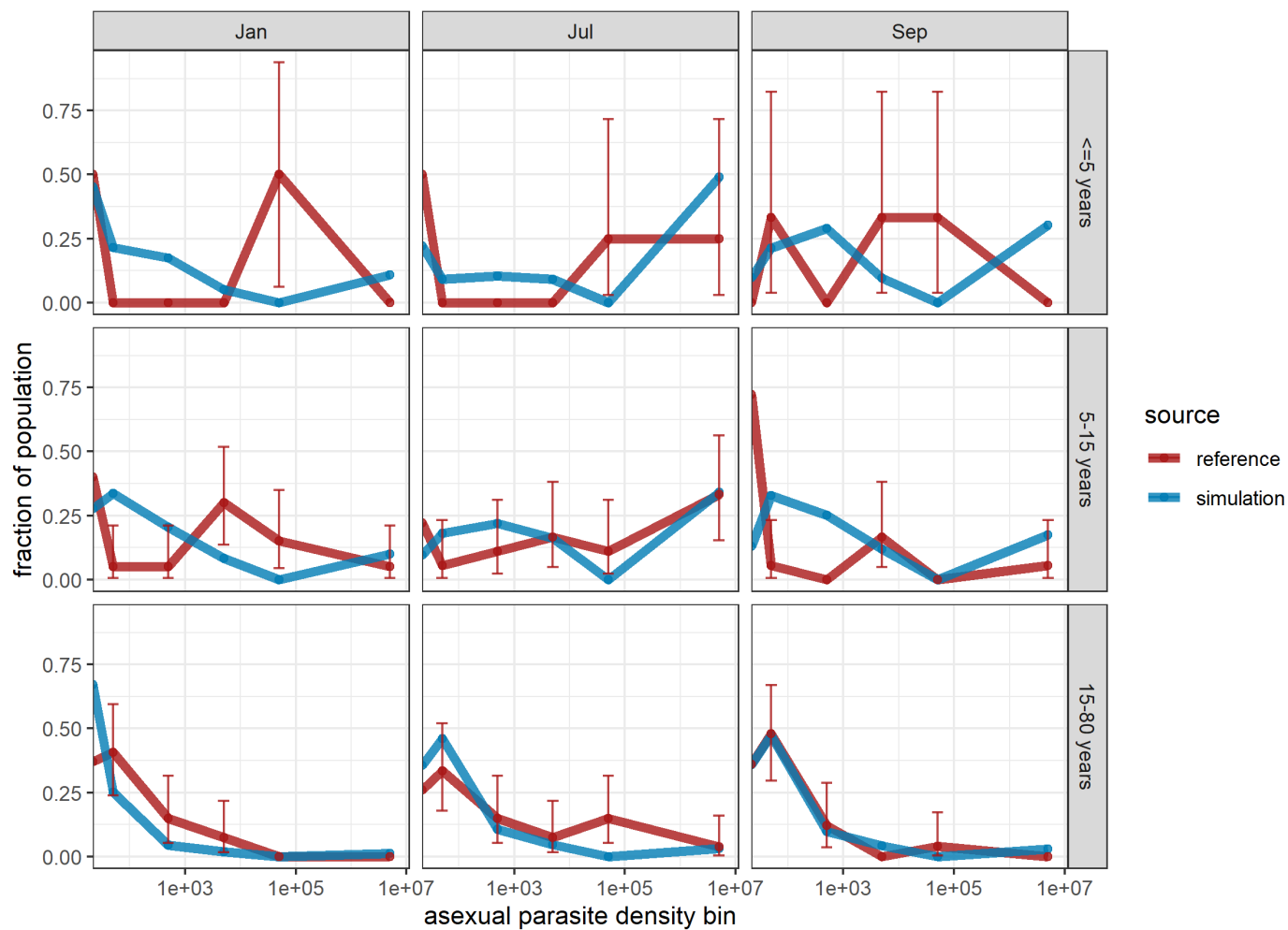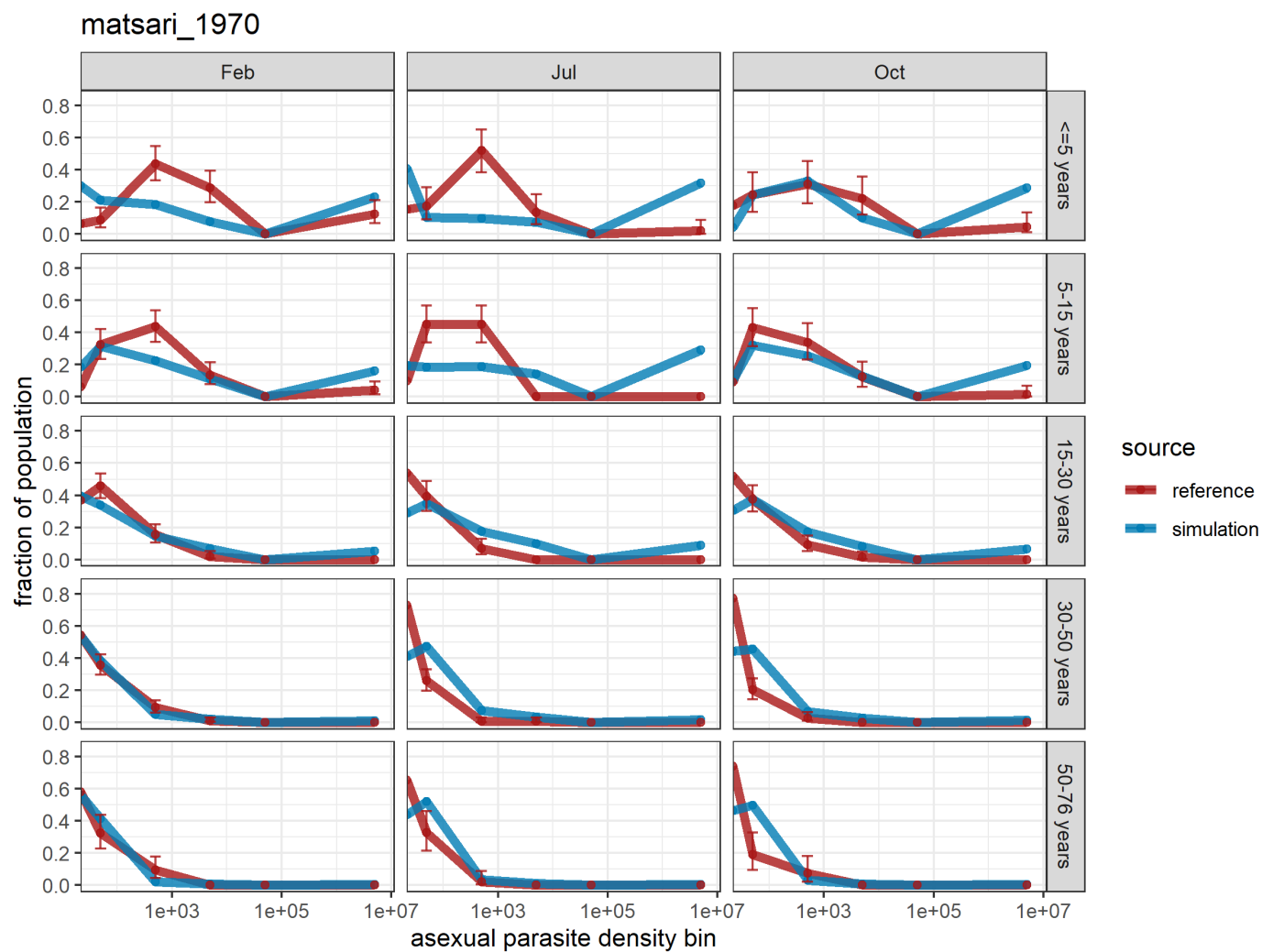
3.6 Asexual parasite density by age

The plots below compare the distribution of parasite densities across ages and seasons from reference datasets and matched simulations. Each plot panel corresponds to a site. Note that some of the reference datasets have small sample sizes, especially in the youngest age groups.

dapelogo_2007

laye_2007

Malaria Model Validation Results

matsari_1970

rafin_marke_1970

Malaria Model Validation Results

sugungum_1970

3.7 Gametocyte density by age

The plots below compare the distribution of gametocyte densities across ages and seasons from reference datasets and matched simulations. Each plot panel corresponds to a site. Note that some of the reference datasets have small sample sizes, especially in the youngest age groups.
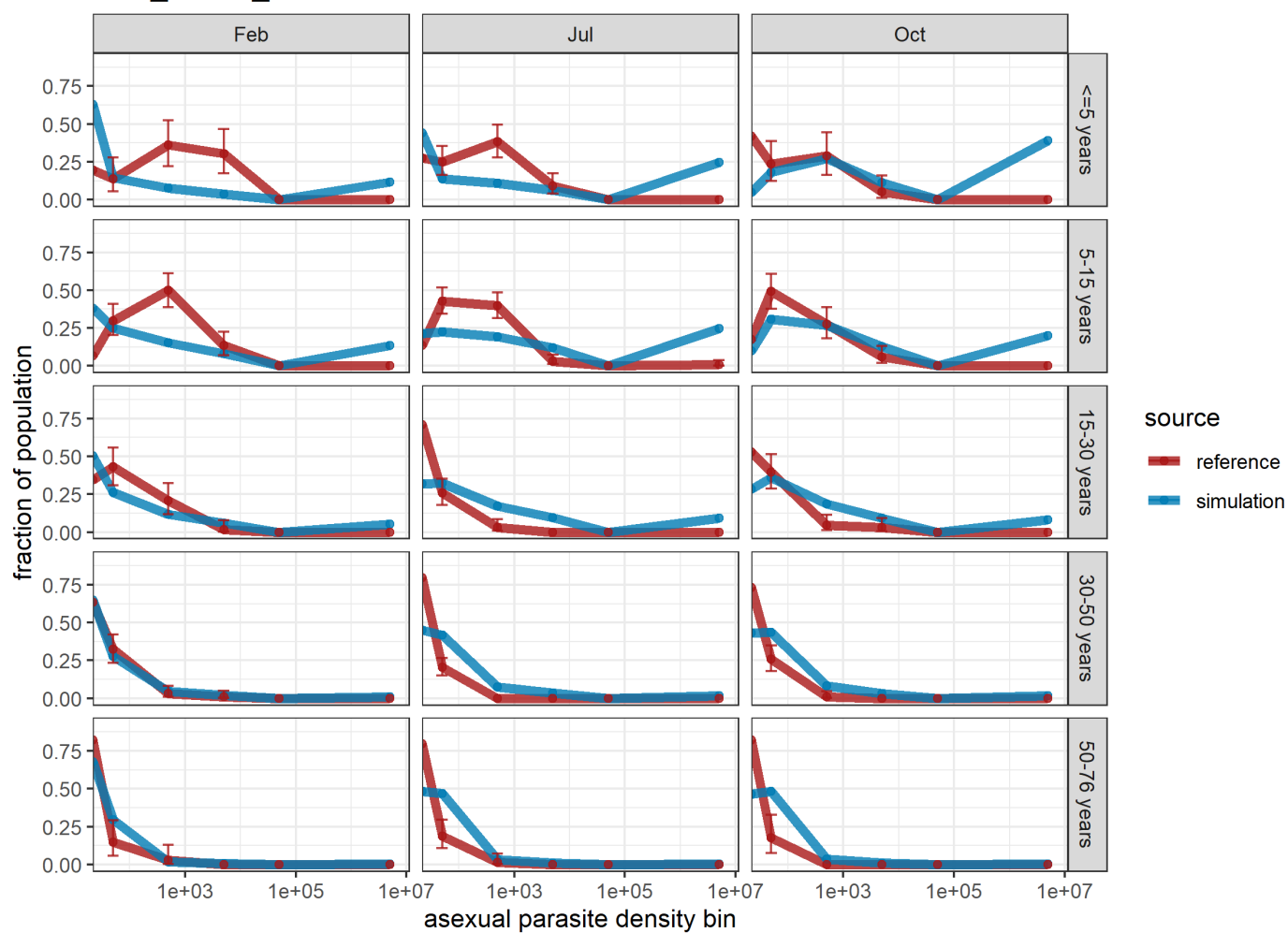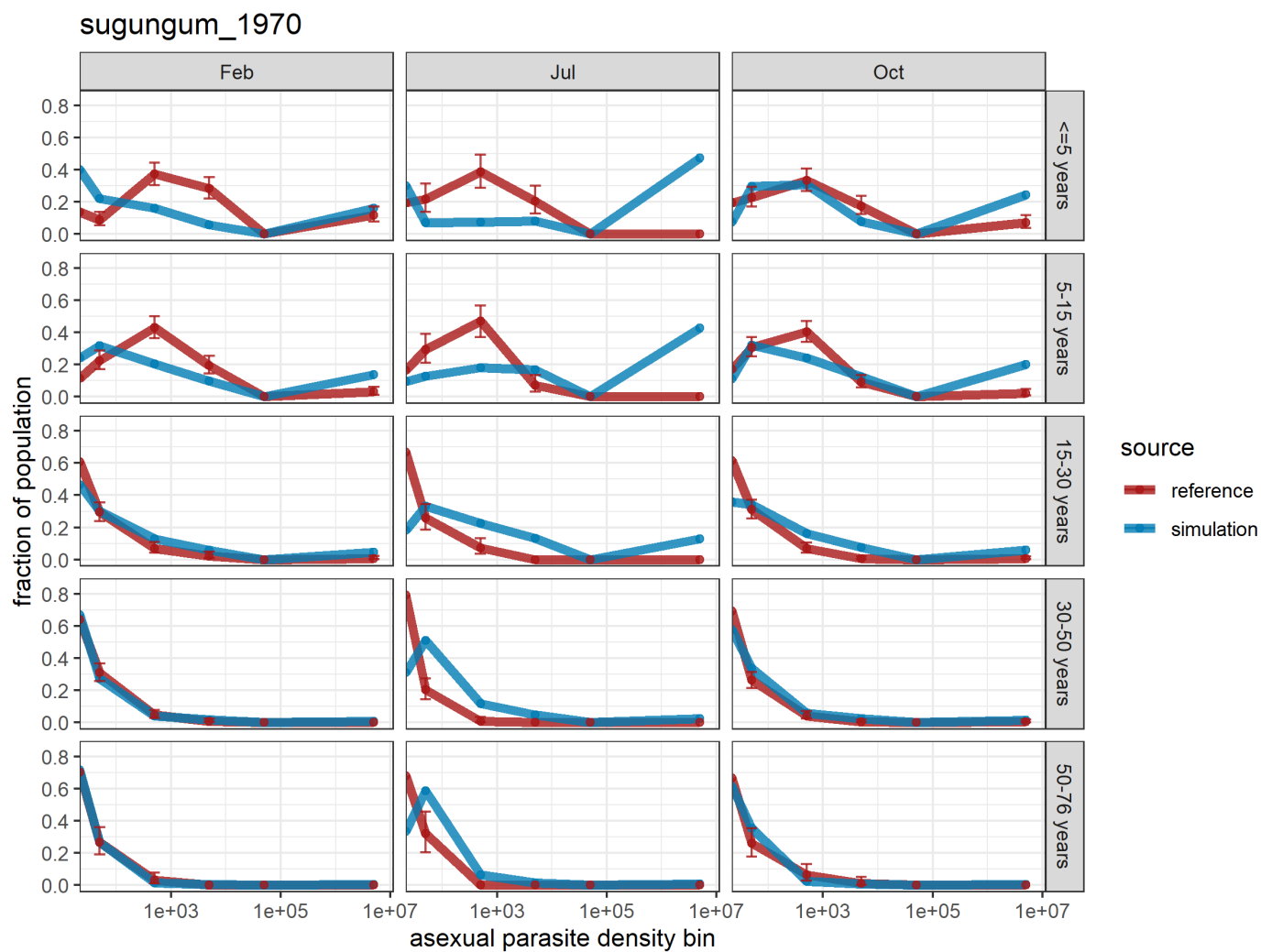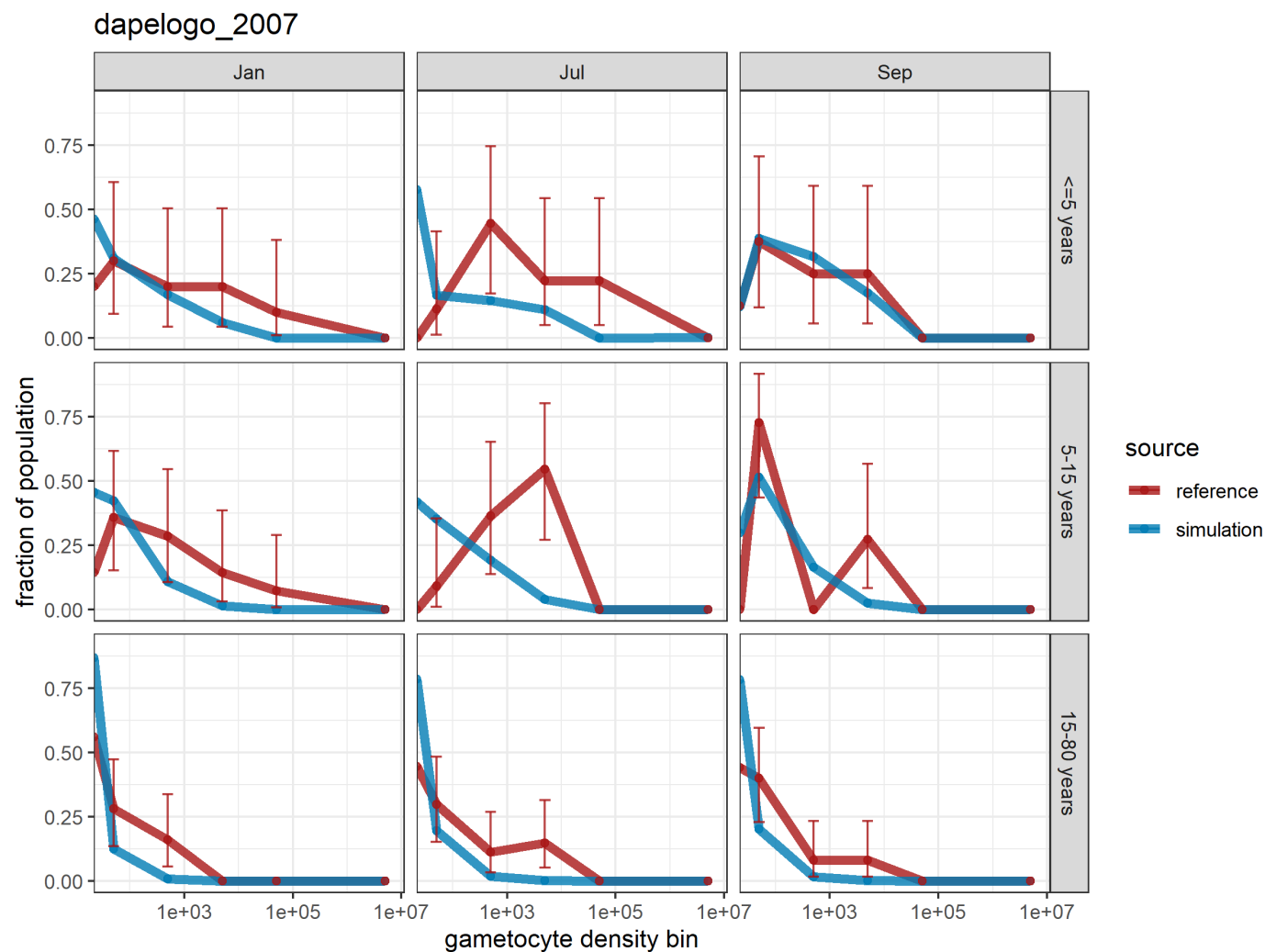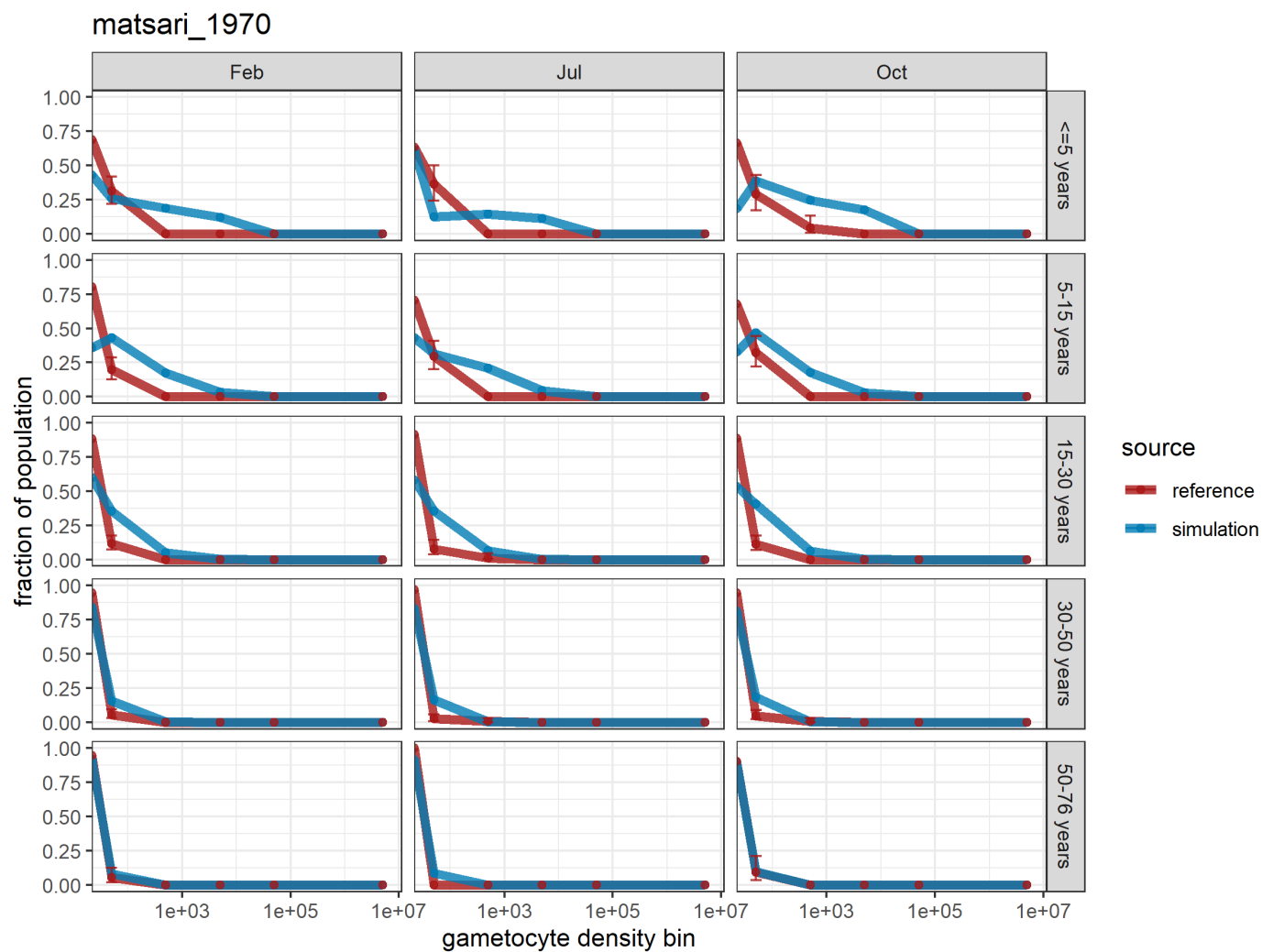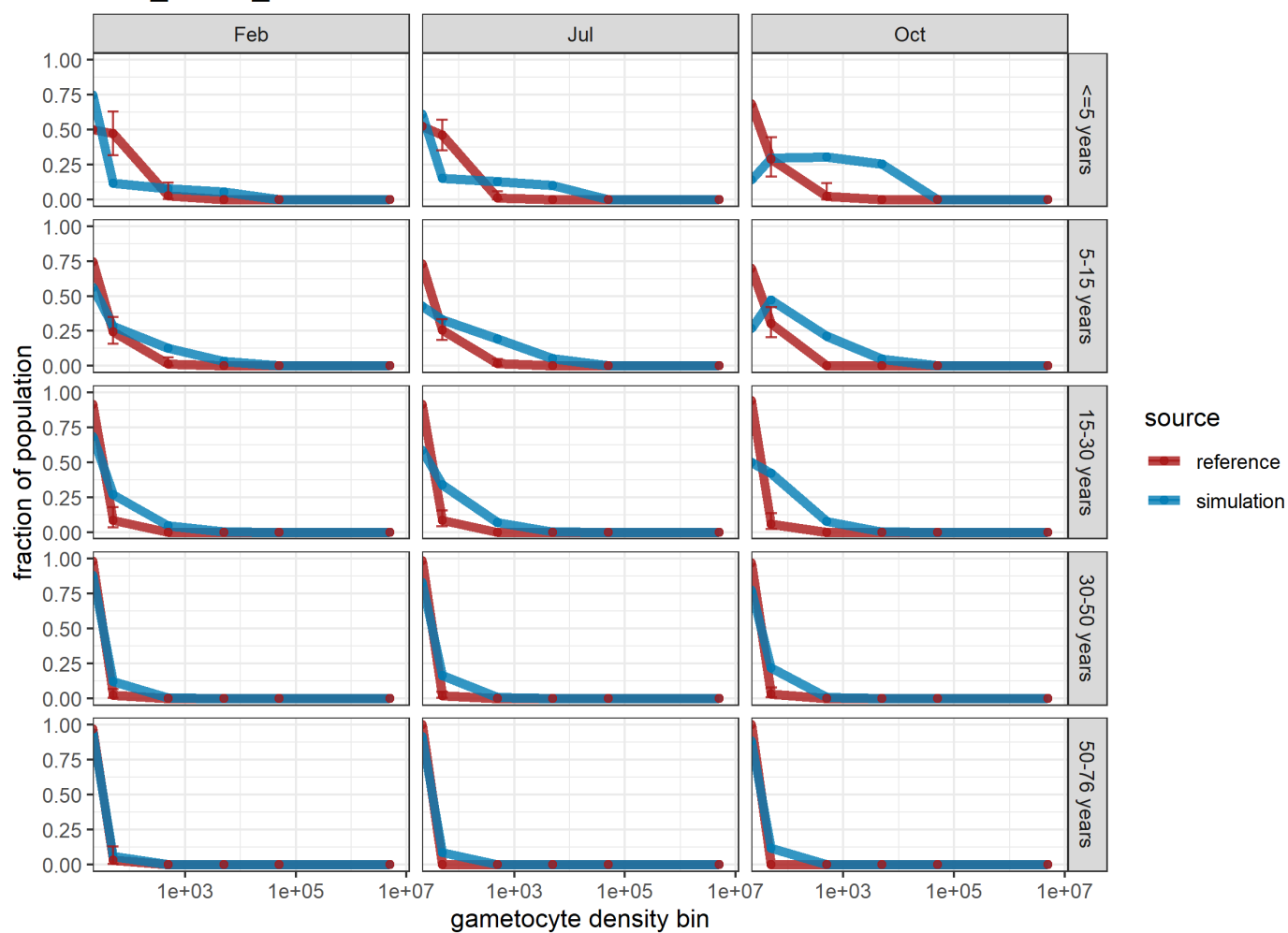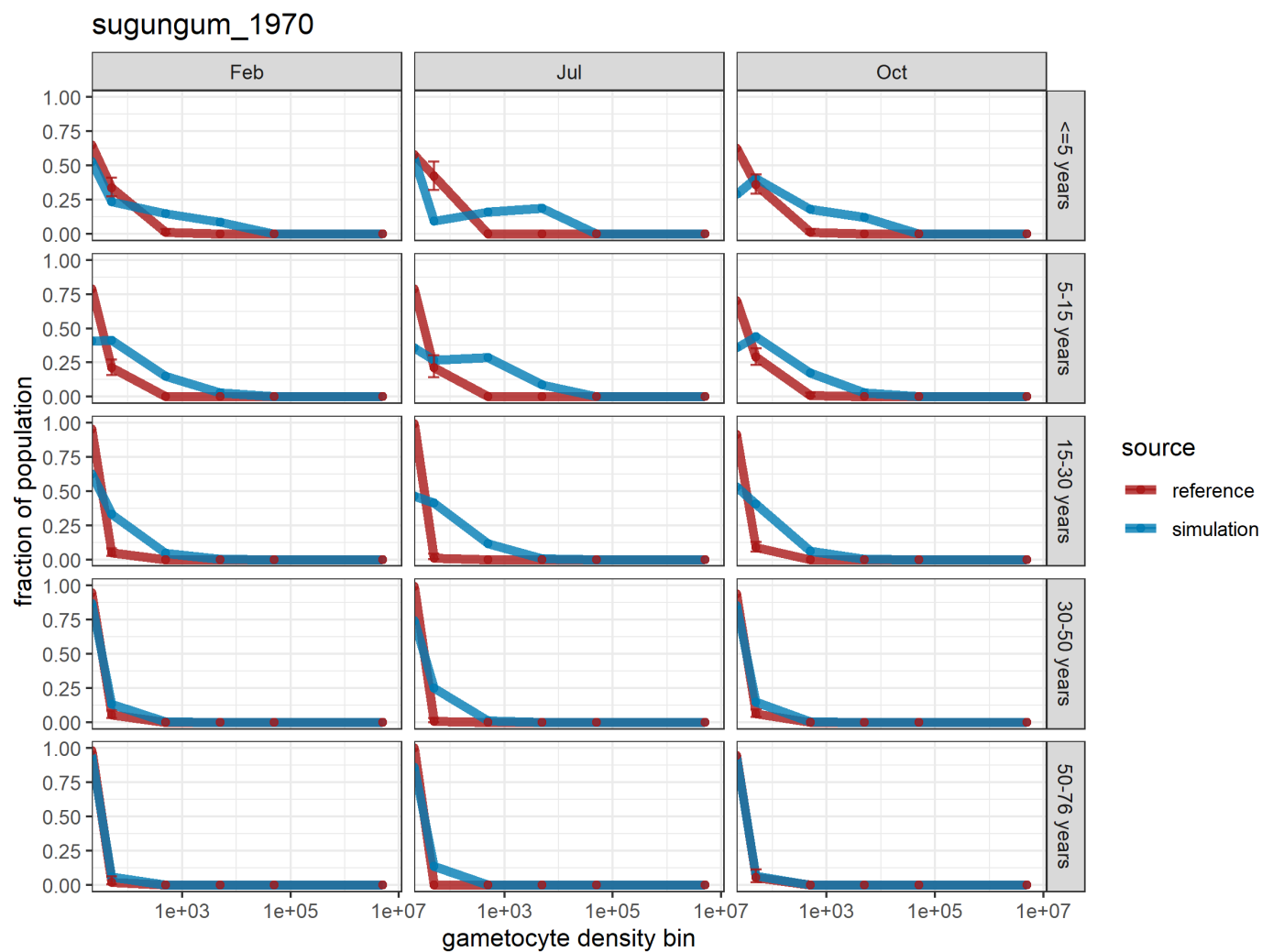
dapelogo_2007

laye_2007

Malaria Model Validation Results

matsari_1970

**Malaria Model Validation Results**
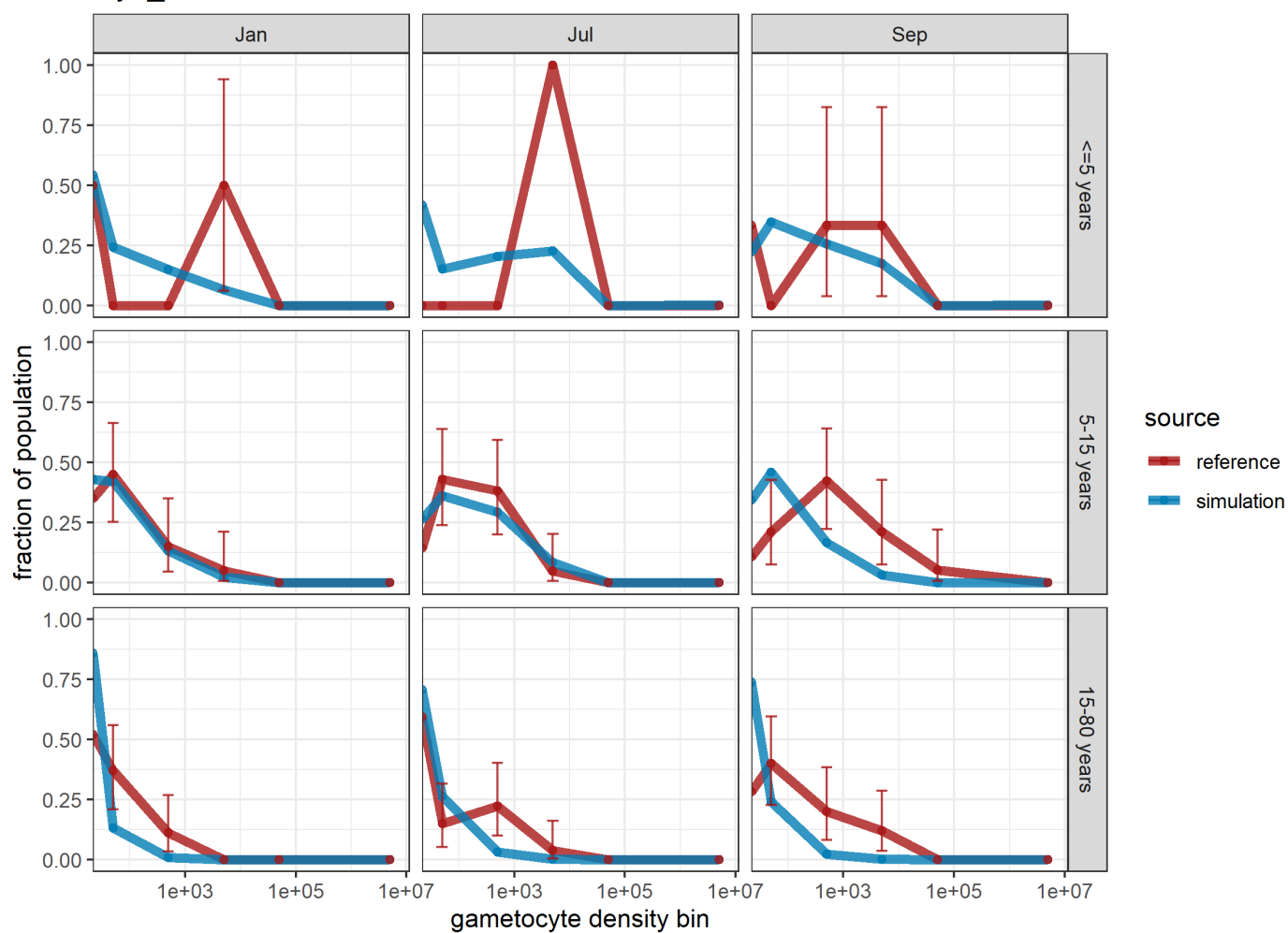
rafin_marke_1970

sugungum_1970

4. Additional comparisons of reference data and matched simulations

4.1 Incidence by age

Below, the plots show the correlation between the simulation and reference incidence in each age group (left plot) and the correlation between the simulation and reference slopes when moving between the incidence in one age group and the incidence in the next oldest age group (right plot).
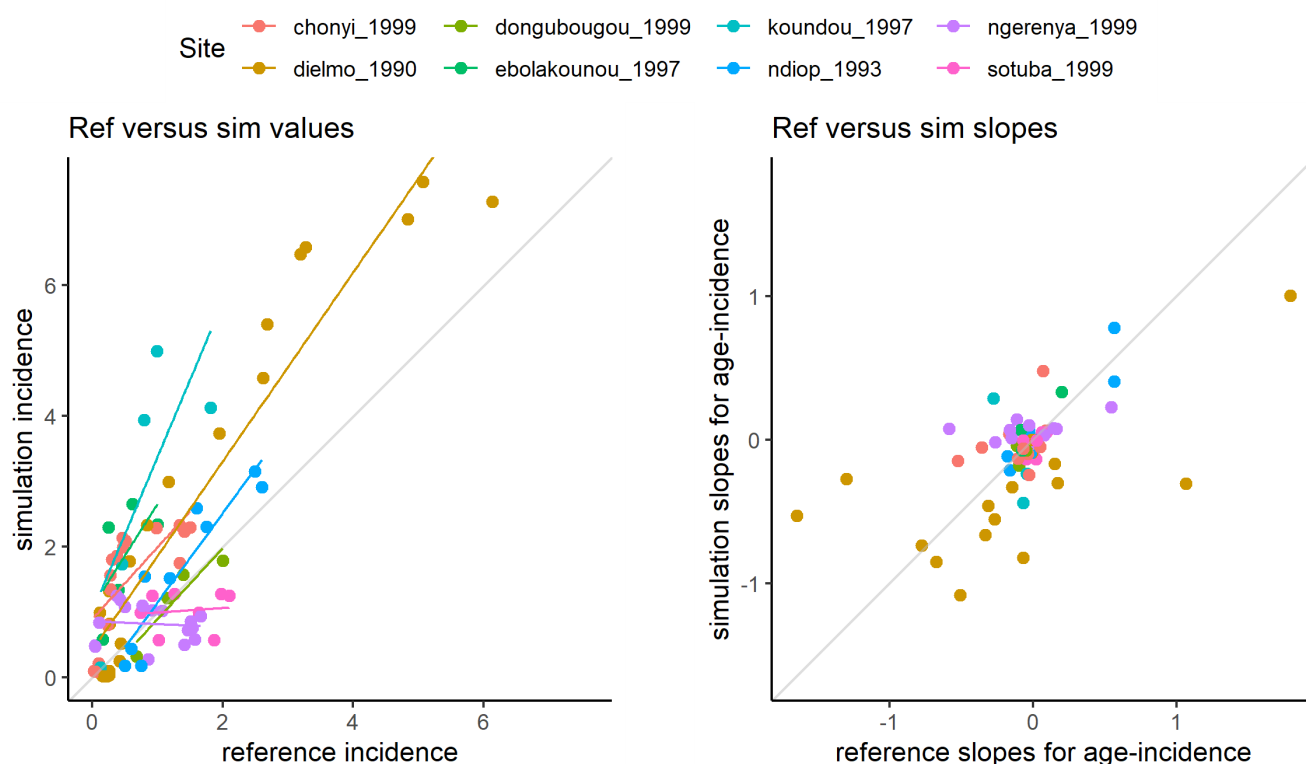
Site	mean_rel_diff	mean_abs_diff	corr_slope	corr_r_squared	mean_rel_slope_diff	mean_abs_slpe_diff
chonyi_1999	2.35	0.95	1.1	0.51	7.28	0.13
dielmo_1990	1.41	1.24	1.45	0.91	2.66	0.37
dongubougou_1999	0.2	0.2	1.09	0.86	0.66	0.06
ebolakounou_1997	3.53	1.35	1.58	0.38	5.35	0.08
koundou_1997	2.45	2.15	2.37	0.58	3.97	0.29
ndiop_1993	0.47	0.52	1.37	0.86	1.47	0.08
ngerenya_1999	2.17	0.61	-0.04	0.01	2.95	0.16
sotuba_1999	0.37	0.57	0.07	0.01	2.04	0.07

4.2 Prevalence by age

Below, the plots show the correlation between the simulation and reference prevalence in each age group (left plot) and the correlation between the simulation and reference slopes when moving between the prevalence in one age group and the prevalence in the next oldest age group (right plot).
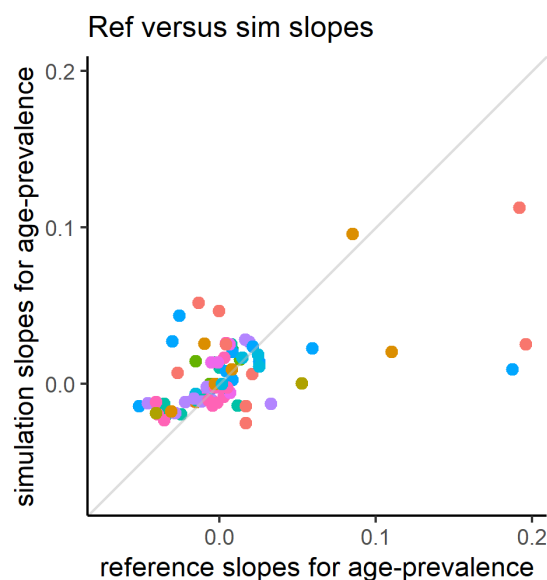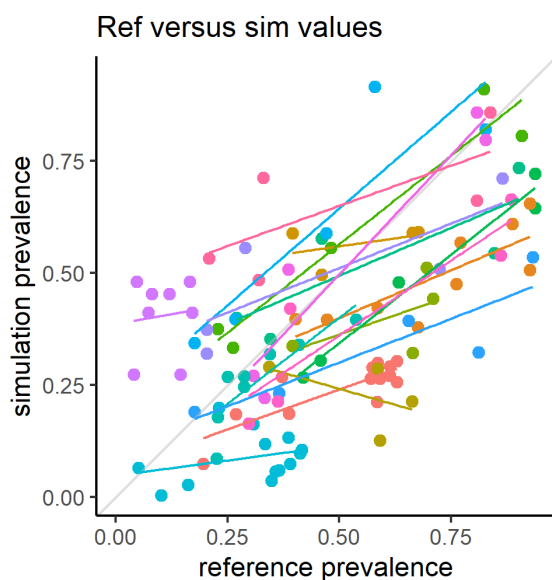

Site	mean_rel_diff	mean_abs_diff	corr_slope	corr_r_squared	mean_rel_slope_diff	mean_abs_slpe_diff
chonyi_1999_month7	nan	nan	0.36	0.68	25.3	0.05
dielmo_1990_monthmultiple	0.29	0.22	0.43	0.66	0.9	0.02
dongubougou_1999_monthmultiple	0.2	0.1	0.14	0.19	0.83	0.03
ebolakounou_1997_monthmultiple	0.53	0.32	-0.29	0.26	1.07	0.01
koundou_1997_monthmultiple	0.33	0.21	0.35	0.33	2.55	0.01
matsari_1970_month10	0.26	0.1	0.79	0.93	1.28	0.01
matsari_1970_month2	0.3	0.19	0.79	0.97	inf	0.0
matsari_1970_month7	0.26	0.14	0.43	0.67	1.2	0.01
ndiop_1993_monthmultiple	nan	nan	0.77	0.72	0.5	0.01
ngerenya_1999_mo						

nth7	nan	nan	0.14	0.09	1.07	0.04
rafin_marke_1970_month10	0.45	0.15	0.86	0.78	0.55	0.01
rafin_marke_1970_month2	0.37	0.26	0.39	0.79	0.62	0.01
rafin_marke_1970_month7	0.56	0.18	0.4	0.65	51.82	0.01
sotuba_1999_monthmultiple	3.87	0.3	0.18	0.01	4.39	0.01
sugungum_1970_month10	0.17	0.07	1.07	0.91	2.36	0.01
sugungum_1970_month2	0.31	0.17	0.67	0.83	2.88	0.01
sugungum_1970_month7	0.69	0.21	0.36	0.52	2.06	0.02

This page left unintentionally blank