

1 A hierarchical Bayesian quantitative microbiological risk
2 assessment model for *Salmonella* in the sheep meat food chain

3 Thomas Rawson^{*a}

4 ^a*Mathematical Ecology Research Group, University of Oxford, Department of Zoology,*
5 *Oxford, OX1 3PS, U.K.*

6 ^a*thomas.rawson@zoo.ox.ac.uk*

7 January 6, 2022

^{*}Present correspondence: Department of Infectious Disease, School of Medicine, St Mary's Hospital, Praed Street, London W2 1NY. email: t.rawson@imperial.ac.uk

Multiple foodborne routes of *Salmonella* infection have been observed; however, the majority of the literature to date has been dominated by research into the most frequently observed reservoirs, such as chicken, beef, and pork. While less commonly observed, outbreaks of *Salmonella* within sheep meat still occur, requiring extensive investigation by food safety inspectors. Risk assessment models inform policy makers and investigators of the risks posed by pathogens at each stage of the food chain, and help suggest at which stages in the food chain outbreaks are likely induced. This work is the first risk assessment into the prevalence of *Salmonella* throughout the sheep meat food chain, from farm to fork. A Bayesian evidence-synthesis model is used, informed by data gathered from 27 individual studies - an exhaustive search of the existing literature, to express and enumerate the current understanding of *Salmonella* prevalence in the sheep meat food chain in the form of probabilities of colonisation throughout the food chain. The resulting posterior estimate projects that 9 (0 - 29 95% HDI) UK individuals are likely to fall ill with salmonellosis due to sheep meat every year. A variance-based sensitivity analysis reveals that the abattoir module is the stage of greatest bacterial proliferation, highlighting it as the most probable source of outbreaks, though not to the exclusion of other factors.

24

Keywords: *Salmonella*; Risk Assessment; Mathematical Modelling; Bayesian Statistics; Food Microbiology

26

1 INTRODUCTION

Reported incidents of human salmonellosis in the UK have varied from 7,000 to 12,000 per year between 2007 and 2016 (Public Health England (PHE), 2018). While it is difficult to accurately trace the cause of most clinical cases, Smerdon et al. (2001) estimate that roughly 16% of cases are likely caused by the consumption of contaminated red meat, with beef and pork being the most frequently implicated in the UK. As a result of this, the vast majority of the literature investigating *Salmonella* in the food chain has predominantly focused on the pork, beef, and chicken chains, with little attention given to lamb and mutton. While sheep meat is less commonly eaten in the UK and the EU (Organisation for Economic Cooperation and Development (OECD), 2019) compared to other meats, it has additionally been found that sheep meat is less likely to be colonised by pathogenic bacteria at the point of sale than other meat products (Food Standards Agency (FSA), 2010; Busani et al., 2005; Davies et al., 2004; Little et al., 2008). Despite this, there have nonetheless been rare reports of larger *Salmonella* outbreaks linked to the consumption of sheep meat (Hjartardóttir et al., 2002; Evans et al., 1999; Synnott et al., 1993; Carson and Davies, 2018). The dearth of data and investigative study into the spread of *Salmonella* within the lamb and mutton food chain to date means that food safety inspectors must engage in lengthy and wide-reaching investigations across each stage of the food chain to identify the root cause when such outbreaks do occur, uninformed by any lessons that can be learned from previous outbreaks.

As such, this project has utilised the varied existing global studies on *Salmonella* colonisation of sheep, and incorporated these to inform a Bayesian quantitative microbiological risk assessment (QMRA) model, to identify the points in the food chain at which *Salmonella* outbreaks are most likely to propagate, and to help determine the risk to human health posed by the consumption of sheep meat. Bayesian statistical techniques are well-established as an effective method for compiling risk assessments (Beaudequin et al., 2015; Albert et al., 2008), as the implicit uncertainty in parameter estimation is transparently conveyed in the resulting posterior distributions. More importantly, hierarchical Bayesian networks such as those employed in this project perform far better than their frequentist counterparts when challenged with limited (and missing) data (Dorazio, 2016; Verdonck et al., 2001), and their prescriptive nature facilitates collaborative discourse regarding initial model formulation (Efron, 2005). Since *Salmonella* outbreaks via sheep meat are rare, available relevant data is sparse, thus a Bayesian methodology ensures that the uncertainty and limitations of any such findings are appropriately displayed in the resulting probability distributions.

The model presented simulates a full ‘farm-to-fork’ network, utilising data describing on-farm disease prevalence, transportation of livestock, slaughtering procedure, consumer purchase habit, cross-contamination, and the dose-response relationship. This work is compiled to provide policy makers and food safety inspectors with a probabilistic summary of the risks posed by *Salmonella* at each point in the production

line of sheep meat products, as informed by the current research landscape. Furthermore, the model is designed to be module-based, to allow it to be later used or expanded in response to active investigation under known outbreak-specific parameters, or expansion of the knowledge set.

2 METHODS

An introductory guide to the methodology underpinning Bayesian statistics is provided in Appendix 1.

The model presented separates the sheep meat food chain into six dependent modules: on-farm, meat processing, retail, hygiene, exposure and illness. Each module comprises of a number of variables of interest, designed around the data sets available, linked through a series of deterministic and probabilistic dependencies. Fig. 1 below presents a graphical display of the model structure, while Table 1 defines and describes all of the included variables and parameters. Following this, an in-depth explanation of each module formulation is provided below.

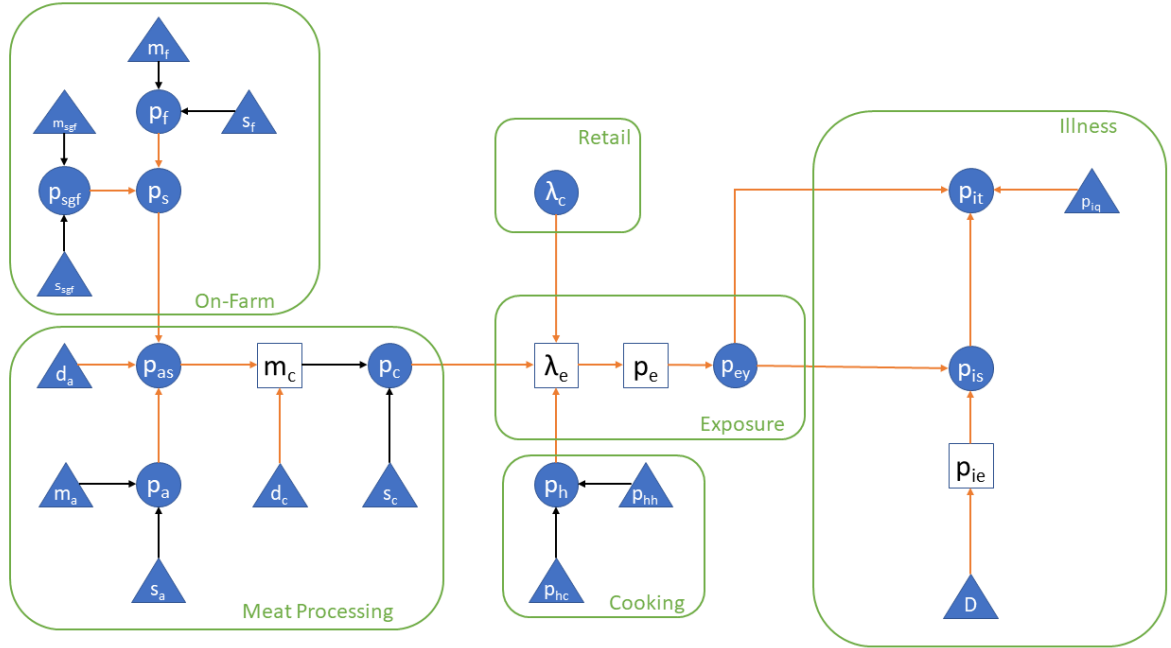


Fig. 1. Model framework. The compartmental network of dependencies of the model. Each node represents a model variable. Circles indicate specific variables of interest. Triangles represent distribution defining parameters. Squares indicate deterministic placeholder variables. All variable definitions and distributions are listed in Table 1. Orange arrows indicate a purely deterministic dependence, while black arrows indicate a probabilistic dependence. Green boxes indicate the distinct modules of the food chain.

2.1 Model Structure

Throughout the model, the notation p_x indicates a probability, while m_x and s_x are hyperparameters, the mean and standard deviation respectively, defining the hierarchical normal distribution from which associated p_x parameters are drawn.

Variable	Definition	Distribution
On-farm		
p_f	Probability a flock is colonised.	$\text{logit}(p_f) \sim N(m_f, s_f)$
m_f	Mean value of the logit probability of flock colonisation.	$m_f \sim N(-0.5, 20)^*$
s_f	Standard deviation value of the logit of flock colonisation probability.	$s_f \sim \text{Uniform}(0, 50)^*$
p_{sgf}	Probability a sheep in a colonised flock is, itself, colonised.	$\text{logit}(p_{sgf}) \sim N(m_{sgf}, s_{sgf})$
m_{sgf}	Mean value of the logit probability of sheep (given colonised flock) colonisation.	$m_{sgf} \sim N(-2, 1.5)^*$
s_{sgf}	SD value of the logit of sheep (given colonised flock) colonisation probability.	$s_{sgf} \sim \text{Uniform}(0, 20)^*$
p_s	Probability that any sheep is colonised.	$p_s = p_f p_{sgf}$
Meat processing		
p_a	Probability a sheep becomes colonised by one night in lairage.	$\text{logit}(p_a) \sim N(m_a, s_a)$
m_a	Mean value of the logit probability of p_a .	$m_a \sim N(-2, 1)^*$
s_a	Standard deviation value of the logit probability of p_a .	$s_a \sim \text{Uniform}(0, 10)^*$
p_{as}	Probability that a sheep is colonised immediately before slaughter.	$p_{as} = 1 - e^{-(p_s + d_a p_a (1 - p_s))}$
d_a	Days a sheep spends in lairage.	$d_a \sim N(0.7, 1) > 0^\dagger$
p_c	Probability that a sheep carcass is colonised.	$\text{logit}(p_c) \sim N(m_c, s_c)$
m_c	Mean value of the logit probability that a sheep carcass is colonised.	$m_c = \text{logit}(p_{as}) d_c$
s_c	Standard deviation value of the logit probability of p_c .	$s_c \sim \text{Uniform}(0, 20)^*$
d_c	Carcass processing adjustment parameter.	$d_c \sim N(0.7, 1.5)^*$
Cooking		
p_{hc}	Probability of <i>Salmonella</i> transfer to other surface from meat.	$p_{hc} \sim \text{Beta}(7, 33)^1$
p_{hh}	Probability of poor meat preparation hygiene	$p_{hh} \sim \text{Beta}(8, 28)^2$
p_h	Probability of contamination from colonised meat	$p_h = p_{hh} p_{hc}$
Retail		
λ_c	Intensity of sheep meat purchases in a four-day period.	$\lambda_c \sim \text{Gamma}(1, 1)^*$
Exposure		
λ_e	Intensity of <i>Salmonella</i> exposure via sheep meat in a four-day period.	$\lambda_e = p_c p_h \lambda_c$
p_e	Probability of <i>Salmonella</i> exposure via sheep meat in a four-day period.	$p_e = 1 - e^{-\lambda_e}$
p_{ey}	Probability of <i>Salmonella</i> exposure via sheep meat in a one-year period.	$p_{ey} = 1 - (1 - p_e)^{91.25}$
Illness		
p_{is}	Probability of salmonellosis via sheep meat in a one-year period.	$p_{is} = p_{ey} p_{ie}$
p_{ie}	Probability of illness given exposure to <i>Salmonella</i> .	$p_{ie} = 1 - \left(1 + \frac{100D}{51.45}\right)^{-0.1324^3}$
D	Concentration of <i>Salmonella</i> bacteria (CFU/g) within consumed product.	$D \sim \text{Lognormal}(7, 1)^4$
p_{it}	Probability of salmonellosis by any cause.	$p_{it} = \frac{p_{is}}{1 - (1 - p_{iq})(1 - p_{ey})}$
p_{iq}	Population attributable risk (PAR) of sheep for salmonellosis.	$p_{iq} \sim \text{Beta}(1, 2000)^*$

* Wide, uninformative priors.

[†] Prior chosen based on expert consensus from consulted food safety investigators.

¹Smid et al. (2013)

²Albert et al. (2008); Christensen et al. (2001); Yang et al. (1998)

³The World Health Organization (WHO) (2002)

⁴Kilsby and Pugh (1981); Haas (1983)

Table 1. Definitions of the model variables and their associated distributions or deterministic relationships. Model equations are detailed further in section 2.1

81 The first module in the model, on-farm, considers the colonisation status of sheep before they have been
 82 transported for slaughter. Three specific probabilities are included within this module; the probability
 83 that a flock is colonised (p_f), the probability that a sheep is colonised, given that its flock is colonised
 84 (p_{sgf}), and the probability that any sheep is colonised (p_s). Here, a flock is considered colonised if any
 85 sheep in the flock is colonised. These three separate variables are considered as they reflect the different
 86 variety of data that were found when searching the literature. As reported in section 2.2 below, some
 87 studies explore only inter-flock prevalence, while others will study specific animals within a flock.

88

89 p_s can be defined deterministically as the product of p_f and p_{sgf} , as summarised in Table 1, while
 90 p_f and p_{sgf} are themselves inferred from a probability distribution. Due to having multiple data sets
 91 from which to infer the parameter, a hierarchical approach is applied, wherein we model for multiple
 92 such probabilities, unique to each study. As such, we are interested in the distribution from which these
 93 probabilities are drawn. As such we have m_f , the mean probability of flock colonisation, and s_f the stan-
 94 dard deviation of flock colonisations, such that these multiple probabilities of flock colonisation are drawn
 95 from the normal distribution $N(m_f, s_f)$. Specifically, the logit function of the probability is drawn from
 96 this normal distribution, where $\text{logit}(x) = \log\left(\frac{x}{1-x}\right)$. The logit function provides a way of mapping
 97 this distribution to the $[0,1]$ interval, which all probabilities are bounded by. Such formulation is often
 98 considered in hierarchical Bayesian models (Albert et al., 2008), especially in cases where the dependant
 99 variables are not themselves bounded probabilities. This is discussed further in Appendix 1. Likewise we
 100 define m_{sgf} and s_{sgf} such that the probabilities of $\text{logit}(p_{sgf})$ are drawn from $N(m_{sgf}, s_{sgf})$.

101

102 Since we have no pre-existing opinion or bias on any model hyperparameters, we use wide, noninfor-
 103 mative, priors upon m_f , s_f , m_{sgf} and s_{sgf} as shown in Table 1. One may note the use of uniform
 104 improper priors for s_{sgf} and s_f , a practice suggested by Gelman (2006) for cases of low data and high
 105 uncertainty.

106

107 The meat processing module covers the steps of production between a sheep leaving the farm, up to
 108 a final sheep meat product ready for purchase. It is well addressed in the literature that transport and
 109 abattoir processing stand out as crucial potential points of contamination, including contamination due
 110 to intestinal spillage or poor hygiene during evisceration (Duffy et al., 2001; Arguello et al., 2013). Also
 111 raised is the potential for sheep to ingest *Salmonella* shed by other, colonised, sheep due to the densely-
 112 housed conditions when transported or kept in lairage (Grau and Smith, 1974; Purvis et al., 2005; Small
 113 et al., 2002). As such, we consider two distinct steps within the module, also reflecting the variety of
 114 data available. A step whereby total prevalence can be increased by time spent in lairage, and a step
 115 for prevalence to be either increased or reduced by meat processing protocol. As shown in Fig. 1 and
 116 Table 1, we consider three key variables: the probability that a, previously uncolonised, animal becomes

colonised by *Salmonella* through spending one night in lairage (p_a), the probability that an animal is colonised by *Salmonella* immediately prior to slaughter (p_{as}), and the probability that a final sheep meat product is colonised (p_c). $\text{logit}(p_a)$ is hierarchically drawn from a normal distribution with mean m_a and standard deviation s_a , similar to the variables in the on-farm module. Since p_a is the probability of colonisation caused by just one night spent in lairage, we also introduce a variable d_a , the number of days spent in lairage, to be able to investigate the impact caused by shorter/longer than average housing times. Note from Table 1 that we, logically, prevent d_a from being able to take negative values. These factors combine for our final expression for p_{as} :

$$p_{as} = 1 - e^{-(p_s + d_a p_a (1 - p_s))}.$$

Again, this particular form ensures that p_{as} is bounded between 0 and 1. The exponent explores the two scenarios resulting in contaminated sheep prior to slaughter: either an animal is already colonised before transportation (p_s) or it is not colonised ($1 - p_s$), and may become colonised for each day it spends in lairage ($d_a p_a$). As the exponent tends to 0, p_{as} likewise tends to 0, and as the exponent tends to $-\infty$, p_{as} tends to 1. There are numerous steps, from stunning to chilling, when preparing an animal carcass where microbial contaminants may be either reduced, or increased (Bacon et al., 2000; Arguello et al., 2013). As such, it is necessary to model the distribution of p_c such that it has the potential to be both lesser or greater than p_{as} , the probability of colonisation before carcass processing. No data relating to specific abattoir practices could be sourced from the literature (unique to sheep meat), and as such abattoir practice/s must be simplified as a blackbox module. This is incorporated by, again, drawing $\text{logit}(p_c)$ from a normal distribution defined by mean m_c and standard deviation s_c . It is sensible however to assume that this mean value, m_c , is in some way influenced by the colonisation prevalence before processing. Therefore m_c is modelled such that $m_c = \text{logit}(p_{as})d_c$, where d_c is some parameter that either amplifies or reduces the impact of p_{as} . Importantly we note that, when interpreting data from the literature, carcasses are sampled in a variety of different ways. Carcass swabs are taken as indicative of p_c , whereas intestinal samples taken pre-evisceration (another popular sampling method) are considered indicative of p_{as} , as this would imply that the offending bacteria must have been ingested prior to slaughter.

The retail module contains just one parameter, λ_c , the average number of sheep meat meals consumed in a four-day period by an individual, a time-period chosen to reflect the data available to inform it.

If a contaminated piece of meat is purchased, consumer exposure is then dependent on the method of preparation. Indeed, if adequate kitchen hygiene is observed then the meat can be safely consumed with no risk of infection. As such, within the cooking module, we model the probability of exposure from contaminated meat, p_h , as $p_h = p_{hc}p_{hh}$. Here, p_{hh} is the probability of the consumer exposing themselves to the microbial pathogen that has contaminated the meat, due to inadequate meal preparation hygiene.

151 p_{hc} is the probability that *Salmonella* would successfully cross over from the colonised meat to the new
 152 surface. Therefore, a consumer being exposed is dependent on both p_{hc} and p_{hh} occurring. These two
 153 variables are modelled separately to reflect two different varieties of study that may inform our priors,
 154 but also to enable more specific, question-led, investigations using the model, whereby simulations with
 155 particularly invasive species (or particularly poor kitchen hygiene) may be investigated.

156

157 The three previous modules combine to express the overall probability of exposure. For a person to
 158 be exposed to *Salmonella* via sheep meat they must; consume a sheep meat meal (λ_c), that sheep meat
 159 must be colonised by *Salmonella* (p_c), and the bacteria must transfer from the meat to the individual
 160 (p_h). As such, the average number of times a person is exposed to *Salmonella* via sheep meat in a four-day
 161 period is modelled by $\lambda_e = \lambda_c p_c p_h$, which is then used as an intensity parameter within a Poisson distri-
 162 bution to attain the probability of exposure within a four-day period; $p_e = 1 - e^{-\lambda_e}$. This expression is
 163 derived using the probability mass function of the Poisson distribution: $P(k \text{ events in interval}) = \frac{\lambda^k e^{-\lambda}}{k!}$.
 164 Using this expression,

$$\begin{aligned} p_e &= P(\text{at least one exposure event}) \\ &= 1 - P(\text{zero exposures}) \\ &= 1 - P(k = 0) = 1 - \frac{\lambda_e^0 e^{-\lambda_e}}{0!} \\ &= 1 - e^{-\lambda_e}. \end{aligned}$$

165 Lastly, we rescale this probability from being a probability of exposure in a four-day period, to a one
 166 year period,

$$\begin{aligned} p_{ey} &= P(\text{at least one exposure event in a 365-day period}) \\ &= 1 - P(\text{No exposure events in a 365-day period}) \\ &= 1 - P(\text{No exposure events in a four-day period})^{365/4} \\ &= 1 - (1 - P(\text{at least one exposure event in a four-day period}))^{91.25} \\ &= 1 - (1 - p_e)^{91.25}. \end{aligned}$$

167 This time-period will then correctly align with available data for the final illness module.

168

169 The final module, illness, reflects the fact that upon being exposed to *Salmonella* a person may not
 170 necessarily become ill. The probability of a person becoming ill, assuming they have been exposed to
 171 *Salmonella* (p_{ie}), is directly dependent on the number of bacteria ingested. The World Health Organiza-

tion (WHO) (2002) present an approximate beta-Poisson dose-response formulation, such that

$$p_{ie} = 1 - \left(1 + \frac{d}{\beta}\right)^{-\alpha},$$

where d is the number of *Salmonella* ingested, and α, β are shape parameters. Fitting to multiple *Salmonella* dose-response experimental works, The World Health Organization (WHO) (2002) calculate α and β as equal to 0.1324, 51.45 respectively. While this formulation is an approximation to the “true” beta-Poisson model, Schmidt et al. (2013) present ranges of α and β for which the approximation above is reasonable to use ($\beta > 31.13\alpha$), which the proposed shape parameters adhere to.

Lastly, to calculate the number of *Salmonella* consumed, we model the concentration (CFU/gram) of *Salmonella* within the sheep meat product, D , via a log-normal distribution, as it is conventionally attributed (Kilsby and Pugh, 1981; Haas, 1983), using a prior distribution of Lognormal(7, 1) to capture the wide variety of concentrations reported in the outbreaks cited by The World Health Organization (WHO) (2002). Finally, assuming an average portion of sheep meat to be 100g, we write $d = 100D$.

From this, the probability of a person becoming ill with salmonellosis caused by sheep meat in a year (p_{is}) is the probability of exposure from sheep meat in a year (p_{ey}) multiplied by the probability of falling ill from this exposure (p_{ie}).

It is important to have data informing the end point of our hierarchical model, however salmonellosis is so rarely attributed to sheep meat that there is no appropriate data with which to fit to p_{is} . Because of this, we must also consider the probability of contracting *Salmonella* from any source, p_{it} , from which one may back-infer p_{is} by utilising the population attributable risk (PAR) of salmonellosis via sheep meat, p_{iq} . The PAR is defined as the proportion of illness cases that would be prevented if the respective source was eliminated. While there is no existing calculation of the PAR for sheep meat, Smerdon et al. (2001) estimates that 16% of salmonellosis cases are caused by red meat, and a European Food Safety Authority (EFSA) (2008) report noted that only 2 of the 179 reported *Salmonella* outbreaks due to meat products in 2005 were caused by sheep meat. This information was used to justify the chosen prior for p_{iq} .

The given expression for p_{it} , namely, $p_{it} = \frac{p_{is}}{1 - (1 - p_{iq})(1 - p_{ey})}$, warrants further explanation. Consider the slight reformulation of Bayes’ Theorem, $P(A) = \frac{P(A|B)P(B)}{P(B|A)}$, and specifically consider the case where the events are defined as $A = \text{‘ill with salmonellosis from any source’}$, and $B = \text{‘ill with salmonellosis from sheep meat’}$. As such, $P(A) = p_{it}$, and $P(B) = p_{is}$. In this case, $P(A|B) = 1$ as, logically, if one is ill due to sheep meat, then you are automatically also ill from ‘any source’. As such, our expression for p_{it} will now hold if $P(B|A) = 1 - (1 - p_{iq})(1 - p_{ey})$. Equivalently, if $P(B^C|A) = (1 - p_{iq})(1 - p_{ey})$, the probability that an individual is **not** ill due to sheep meat, given that they are still ill. Indeed, by

definition, $(1 - p_{iq})$ is the percentage of people ill by any source other than sheep meat, given that the source in question is eliminated, i.e., given that an individual was never exposed to contaminated sheep meat. Therefore $(1 - p_{iq})(1 - p_{ey}) = P(B^C|A)$, and our formulation for p_{it} holds true. This formulation is also used in Albert et al. (2008).

2.2 Data utilised

In this section we briefly summarise the experimental data gathered from the literature regarding *Salmonella* in the sheep meat food chain, from which we construct our posterior estimates of model parameters. Except for data on sheep meat purchasing (detailed below), all data was in the form of binomial observation data (positives/trials), hence all data was fitted using binomial likelihood distributions. Appendix 2 explicitly details the observation data provided from individual studies.

Within the on-farm module, 11 data points in total are used relating to the distribution of p_f , provided by Alvseike and Skjerve (2002), Hjartardóttir et al. (2002), Sören et al. (2015), Sandberg et al. (2002), Yang et al. (2014), Vanselow et al. (2007), and Methner and Moog (2018). All provide a number of flocks tested and a number of flocks found to be positive. We use a binomial distribution as our likelihood function for this data.

Data was also available from many of these studies to inform p_{sgf} , the probability that a sheep is colonised given that *Salmonella* is present within the flock. Sandberg et al. (2002) provided relevant data, however while the number of colonised sheep is reported, only a range of the total number of sheep tested is reported, so there is some uncertainty in the exact number of sheep sampled in this study. Multiple experiments were performed within this study on multiple flocks, and a variety of n (number of sheep tested) values were used from the range reported when encoding the data into the model. Yang et al. (2014) provides numerous relevant data points, however the numbers of sheep tested and the exact number is not reported. Instead the paper provides the mean probability and 95% confidence intervals (95CI) for the probability of a sheep being colonised in each flock, provided from a binomial distribution. Specific values were back-computed from the reported mean and intervals, however there will be some margin of error in this process. Hjartardóttir et al. (2002) had relevant data regarding untreated sheep heads, and another experiment tested lambs in the spring. These lamb faecal samples were however pooled, with samples from six lambs combined into single samples that were then tested for *Salmonella*. As such, one can only compute, via a binomial likelihood, the probability that none of the six sheep in a sample is colonised. This can, however, then be transformed into a rough estimate for the probability that an individual sheep is colonised by the relationship $P(\text{pooled sample is negative}) = (1 - P(\text{a sheep is colonised}))^6$. Similarly Vanselow et al. (2007) reports an experiment using pooled samples which is included in the same fashion.

Hjartardóttir et al. (2002), Sandberg et al. (2002), Vanselow et al. (2007), and Esmaeili and Rahmani (2016) all provide information relating to the overall probability that any sheep is colonised, p_s .

Within the meat processing module, Grau and Smith (1974) was the only study found that presented clear experimental details into the impact of contaminated lairage on the housed sheep. This was used to provide information on p_a , the probability that an uncolonised sheep becomes colonised by *Salmonella* due to spending one night in lairage.

Data relating to p_{as} is obtained from Hanlon et al. (2018), Sören et al. (2015), Davies et al. (2004), Bonke et al. (2012), Nottingham and Urselmann (1961), Samuel et al. (1981), Kumar et al. (1973), and Duffy et al. (2010). One data point taken from Kumar et al. (1973) relates specifically to goats, however it was felt that this was relevant enough to still warrant inclusion.

Data relating to the probability that any final carcass or sheep meat product is colonised, p_c , consists of both abattoir-level carcass swabs, and tests of meat samples from supermarket shelves. This data is gathered from Kane (1979), Milnes et al. (2008), Duffy et al. (2001), Duffy et al. (2010), Little et al. (2008), Vanderlinde et al. (1999), Phillips et al. (2001), Busani et al. (2005), Wong et al. (2007), and Food Standards Agency (FSA) (2010).

For the retail module, all retail habit data was provided on request from the National Diet and Nutrition Survey (NDNS). The raw data used is not yet available publicly. 9,425 people kept highly detailed food surveys for a four-day period. From this, information regarding the number of times each individual consumed a sheep meat product in that four-day period was extracted. This information was used to inform the intensity of lamb consumption, λ_c .

Finally, data on the total number of human salmonellosis cases, regardless of cause, was taken from the Public Health England (PHE) (2018) report on salmonellosis. This was used to inform the final variable in the model chain, p_{it} .

Due to the Bayesian modelling approach, variables that do not have associated data to inform them specifically will still be informed by data from further down in the food chain, however the choice of priors is of particular importance for these variables. As seen in Table 1, wide and noninformative priors were placed on our hierarchical parameters, such as m_f , s_f , m_a and s_a , however it is worth mentioning the reasoning behind the priors for the other parameters. For parameter d_a , the number of days a sheep is kept in lairage, a prior of $d_a \sim N(0.7, 1) > 0$ is used. A normal distribution with mean 0.7 and standard deviation 1. Note the rejection of any negative values, as logically a sheep cannot spend a negative amount of time in lairage. This distribution was chosen based on the observations of Purvis et al. (2005)

277 and Department for Environment, Food & Rural Affairs (DEFRA) (2006), which report that most sheep
 278 would be processed the day they arrive at the abattoir, and those that are kept in lairage paddocks would
 279 usually be kept for only one or two nights.

280

281 Within the cooking module, the prior for p_{hc} , the probability of *Salmonella* transfer from colonised
 282 meat to another surface, was based upon the study by Smid et al. (2013), which reported a mean approx-
 283 imate value of 0.19. To capture this, a prior of $p_{hc} \sim \text{Beta}(7, 33)$ is used. Similarly, for the probability
 284 of poor consumer hygiene, p_{hh} , we use the same prior as Albert et al. (2008), who based their prior of
 285 $p_{hh} \sim \text{Beta}(8, 28)$ on the studies of Christensen et al. (2001) and Yang et al. (1998).

286 2.3 Numerical simulation

287 The model was run using JAGS (Plummer, 2007), called within R using the ‘rjags’ package (Plummer
 288 et al., 2016). Four iterative chains were used for an initial burn-in period of 5×10^4 iterations, followed
 289 by 10×10^6 further iterations on each chain, thinned at a proportion of 1 out of 1000, meaning that
 290 only 1 in every 1000 iterations was saved with which to calculate the posterior distribution. This is
 291 done to reduce the autocorrelation of the chains to better represent the full distribution, especially in
 292 areas of low probability density. All code written and used for this project is made freely available at
 293 <https://osf.io/em4nj/>.

294 2.4 Variance-based sensitivity analysis

295 A key desirable outcome from this model would be to highlight which factors in the food chain are most
 296 sensitive to affecting the final probability of human illness due to sheep meat. As such, a variance-based
 297 sensitivity analysis was conducted upon a deterministic decomposition of the above model.

298

299 Assuming we assign input values to the nine variables p_f , p_{sgf} , p_a , d_a , d_c , λ_c , p_{hc} , p_{hh} , and D , the
 300 remaining model variables can then be deterministically defined as

$$p_s = p_f p_{sgf}$$

$$p_{as} = 1 - e^{-(p_s + (d_a p_a (1 - p_s)))}$$

$$\text{logit}(p_c) = \text{logit}(p_{as}) d_c$$

$$p_h = p_{hc} p_{hh}$$

$$\lambda_e = p_c p_h \lambda_c$$

$$p_e = 1 - e^{-\lambda_e}$$

$$p_{ey} = 1 - (1 - p_e)^{91.25}$$

$$p_{ie} = 1 - \left(1 + \frac{100D}{51.45}\right)^{-0.1324}$$

$$p_{is} = p_{ey} p_{ie}$$

One then asks which of the nine input variables has the greatest impact on the value of p_{is} , the probability an individual is ill with salmonellosis in a year. Such a question can be answered by calculating two variance-based sensitivity indices. Denote the input variables as X_i for $i \in \{1, 2, \dots, 9\}$, and denote the output of interest, p_{is} , as Y . If we fix one variable, X_i at a value x_i^* , and draw random values for all other input variables, we will obtain a distribution of output values Y for each random selection of input parameters. This distribution of Y values will have an expected value, $E(Y|X_i = x_i^*)$, for each unique x_i^* . As such, the variance of these expected values across all x_i^* values, $V(E(Y|X_i))$ will be an indicator of how much parameter X_i induces change in the output variable. The first-order sensitivity index, S_i , is defined as

$$S_i = \frac{V(E(Y|X_i))}{V(Y)}.$$

Namely, the difference between the innate variance in the output value when all input values are randomised, and the above mentioned variance when parameter i is fixed. S_i will be a value between 0 and 1. A S_i value of 0 would indicate that the mean value of Y was in no way affected by parameter i , implying that parameter i has no impact on the output value. A S_i value of 1 would indicate that the variance in the mean of Y was entirely affected by the parameter in question, meaning that Y is wholly sensitive to parameter i .

316

The main weakness of the first-order sensitivity index is that it only captures the immediate impact of a parameter on the output of interest, and fails to capture the higher order impact that a parameter may have; displaying how a parameter may impact other model parameters in turn, having more complicated downstream impact on the model output. As such, a more accurate sensitivity metric usually employed, is the total-effect index, S_{T_i} , which instead considers the impact caused by fixing all parameters

322 **except for** parameter i , namely $X_{\sim i}$. S_{T_i} is formally defined as,

$$S_{T_i} = 1 - \frac{V(E(Y|X_{\sim i}))}{V(Y)}.$$

323 Again, this index is bound between 0 and 1, with a low value indicating low importance or sensitivity,
324 and a high value indicating higher sensitivity to parameter i . More information on these indices and
325 examples of their application can be found in Saltelli et al. (2008).

326 3 RESULTS

327 Table 2 shows the mean, median and 95% highest density intervals (HDI) of all the posterior distributions.

328
329 Convergence was considered well-achieved through investigation of the trace plots of the chains, the ef-
330 fective sample size (ESS) and Monte Carlo Standard Error (MCSE) of the variables. The Gelman-Rubin
331 statistic, or ‘shrink factor’, is the most-used metric for convergence, with a value close to 1 signifying
332 effective convergence. Heuristically, any shrink factor below 1.1 is considered by Kruschke (2014) to sig-
333 nify sufficient convergence. The presented model run resulted in a multivariate potential scale reduction
334 factor (mpsrf) of 1.00056. Similarly, Kruschke (2014) suggests an ESS of $> 10,000$ as a suitable threshold
335 to signify a sufficiently explored distribution, all of our variables featured an ESS of at least 38,000 (see
336 Appendix 3), with no signs of autocorrelation visible at the chosen thinning interval of 1,000.

337
338 Figs. 2-6 display the density plots and convergence summaries for five variables of particular inter-
339 est, p_s , p_c , p_{ey} , p_{is} , and p_{it} . Such plots are provided for all variables in Appendix 3. These plots were
340 produced using the ‘diagMCMC’ function provided as additional material with Kruschke (2014).

341
342 To plainly express the findings of Figs. 2-6 and Table 2, the probability of a flock of sheep being colonised
343 (p_f) (to any degree) by *Salmonella* has a median value of 0.11, but strong variation between flocks as seen
344 in the data means that the 95% highest density intervals span approximately 0 to 0.71. At an individual
345 sheep level (p_s), approximately one in every 200 sheep (~ 100 to 1000 sheep 95% HDI) are estimated to
346 be colonised by *Salmonella* on the farm, reflecting that flocks will often see incomplete colonisation by
347 *Salmonella*, with only a portion of sheep on the farm colonised.

348
349 Within the meat processing stage, sheep have a low probability of becoming colonised by *Salmonella*
350 while stored in lairage (median $p_a = 0.072$), and effective processes result in a low final risk of contam-
351 ination. At the end of the processing stage, every sheep meat product (p_c) has a 3.655×10^{-7} median
352 probability of contamination with *Salmonella*, informed by how some studies were unable to find any ev-
353 idence of contamination in their experimental studies of store-bought products (Food Standards Agency

Variable	Mean	Median	Lower 95% HDI	Higher 95% HDI
On-farm				
p_f	0.193	0.114	0.003	0.712
m_f	-1.403	-1.411	-3.197	0.404
s_f	9.383	7.362	1.447	23.324
p_{sgf}	0.062	0.042	0.002	0.183
m_{sgf}	-2.879	-2.880	-3.265	-2.495
s_{sgf}	0.964	0.889	0.339	1.728
p_s	5.265×10^{-3}	4.770×10^{-3}	1.486×10^{-3}	1.017×10^{-2}
Meat processing				
p_a	0.104	0.072	0.015	0.290
m_a	-2.735	-2.704	-3.768	-1.769
s_a	3.011	2.399	0.001	7.759
p_{as}	0.079	0.078	0.048	0.113
d_a	1.171	1.070	0.082	2.437
p_c	5.597×10^{-7}	3.655×10^{-7}	2.950×10^{-9}	1.691×10^{-6}
m_c	-5.075	-5.076	-6.617	-3.473
s_c	11.812	11.458	5.710	19.538
d_c	2.064	2.052	1.386	2.770
Cooking				
p_{hc}	0.164	0.158	0.059	0.277
p_{hh}	0.202	0.196	0.079	0.334
p_h	0.033	0.030	0.007	0.065
Retail				
λ_c	0.209	0.209	0.200	0.218
Exposure				
λ_e	2.963×10^{-9}	2.294×10^{-9}	2.841×10^{-11}	7.854×10^{-9}
p_e	2.963×10^{-9}	2.294×10^{-9}	2.841×10^{-11}	7.854×10^{-9}
p_{ey}	2.704×10^{-7}	2.093×10^{-7}	2.592×10^{-9}	7.167×10^{-7}
Illness				
p_{is}	1.695×10^{-7}	1.318×10^{-7}	1.317×10^{-9}	4.443×10^{-7}
p_{ie}	0.631	0.635	0.532	0.726
D	1702.54	1034.30	28.97	5308.48
p_{it}	1.762×10^{-4}	1.669×10^{-4}	7.657×10^{-5}	2.921×10^{-4}
p_{iq}	9.619×10^{-4}	7.997×10^{-4}	7.798×10^{-6}	2.329×10^{-3}

Table 2. Mean, median and HDI intervals of the posterior distributions for each model variable.

p.s

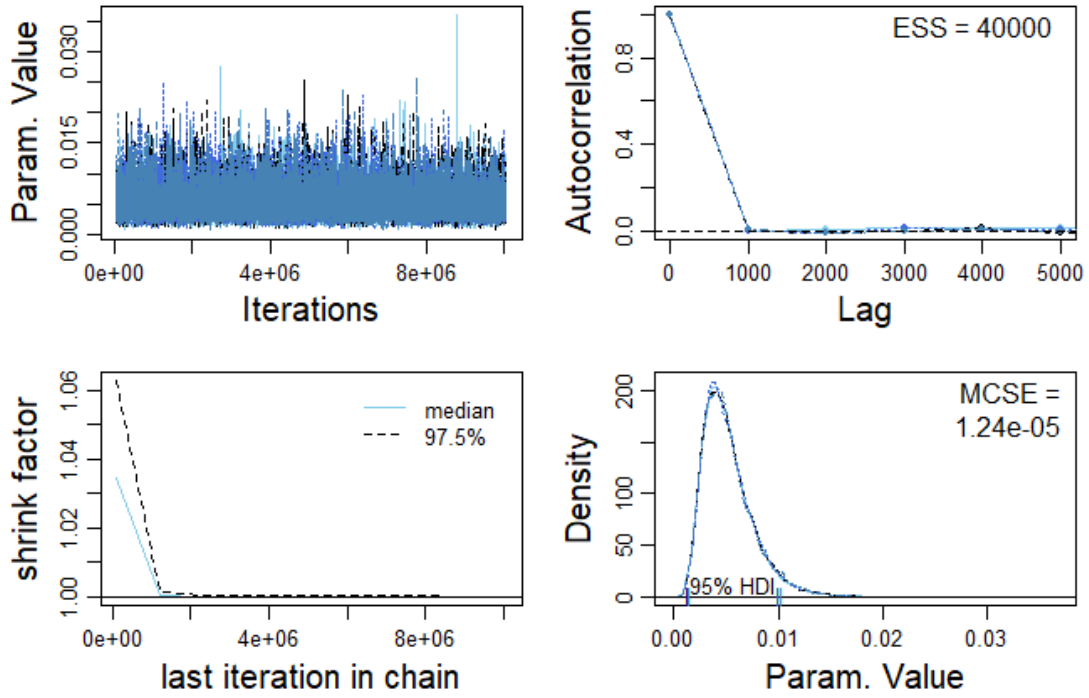


Fig. 2. Posterior of p_s . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.c

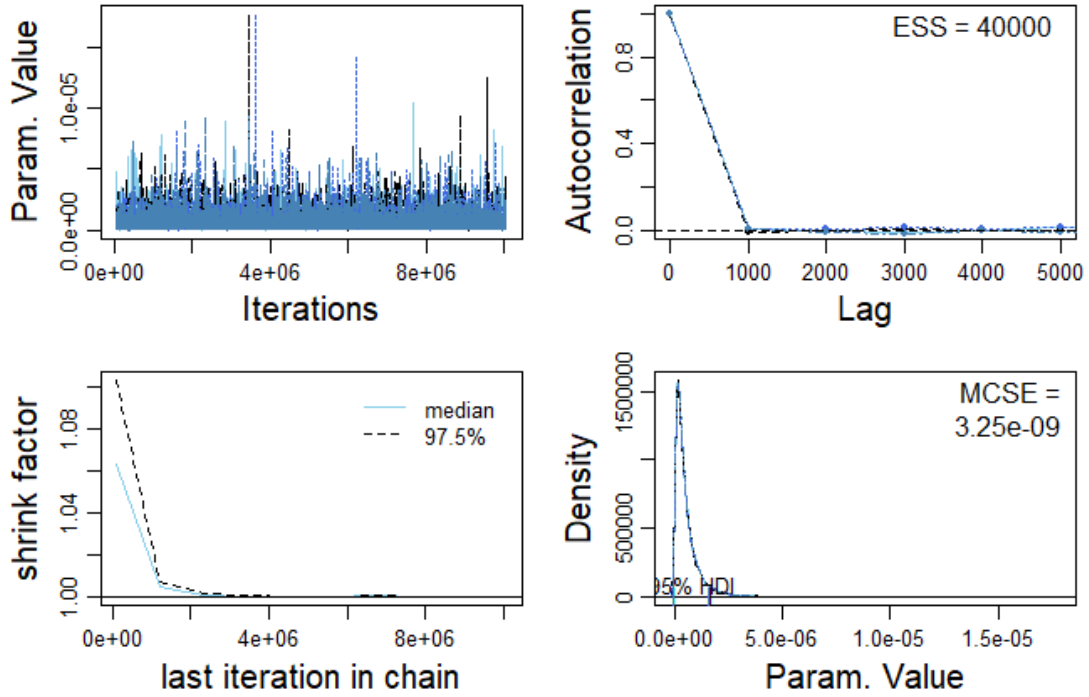


Fig. 3. Posterior of p_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.ey

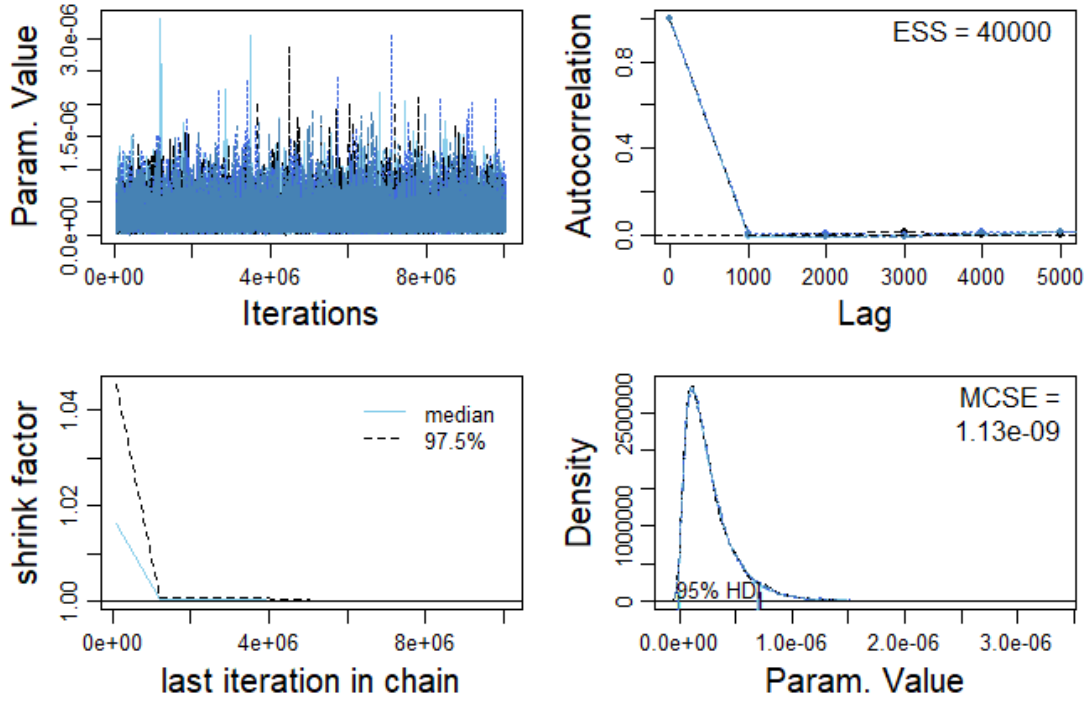


Fig. 4. Posterior of p_{ey} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.is

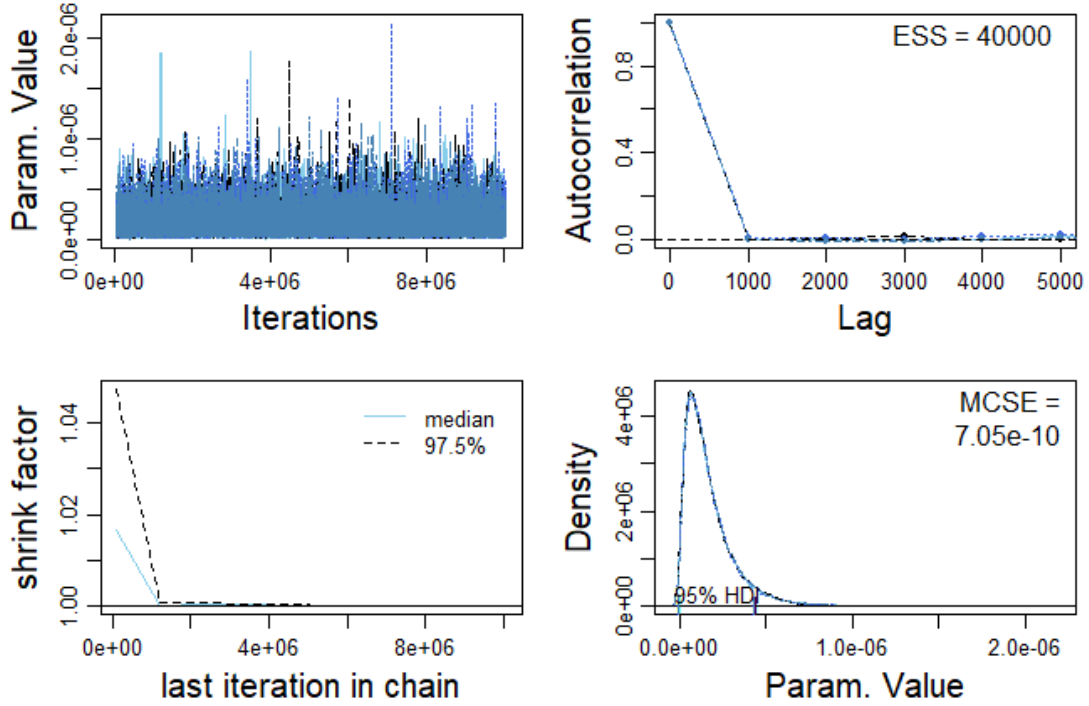


Fig. 5. Posterior of p_{is} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.it

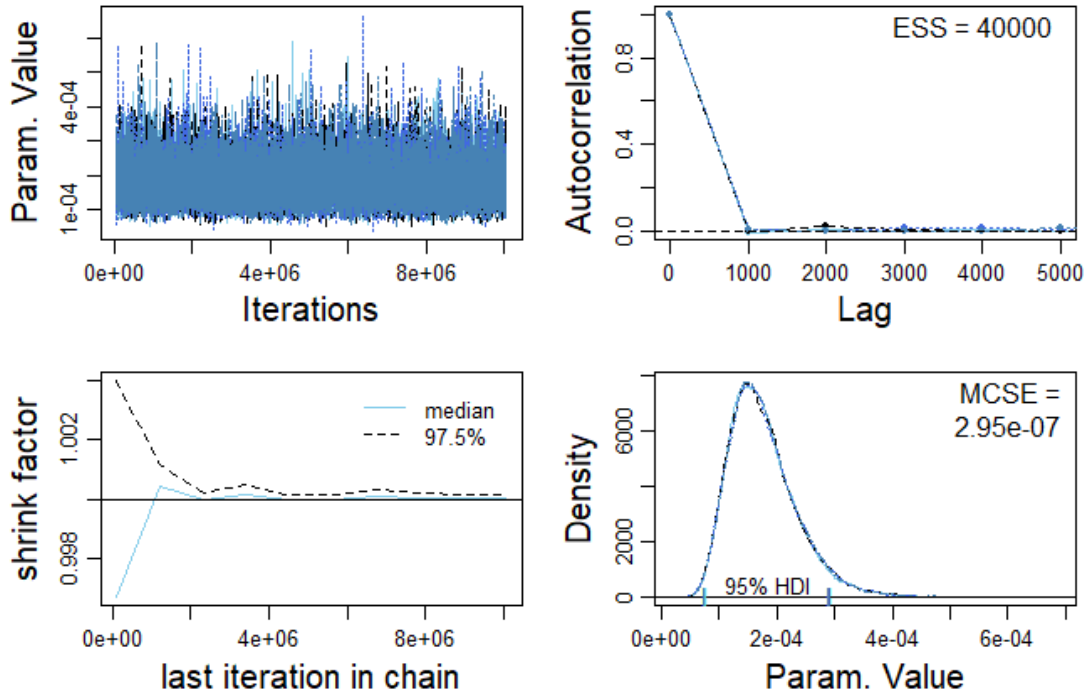


Fig. 6. Posterior of p_{it} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

(FSA) (2010); Duffy et al. (2010); Busani et al. (2005)).

Following a population average of 0.2 meals containing sheep meat every 4 days (λ_c), and a 0.03 probability of consumers ingesting *Salmonella* from contaminated meat (p_h), roughly 14 (0-47 95% HDI) UK consumers are expected to be exposed to *Salmonella* from sheep meat every year (p_{ey} , assuming a UK population of 66 million). Following the illness module, 9 (0 - 29 95% HDI) of these 14 are expected to fall ill with salmonellosis.

The total number of annual expected UK *Salmonella*-caused illnesses by any cause (informed by p_{it} , Fig. 6), is found to be 11,015 (5,054 - 19,297 95% HDI). This equates to 16.69 infections per 100,000 individuals.

3.1 Variance-based Sensitivity Analysis

Table 3 below contains both the first-order and total-effect sensitivity indices for the nine input variables of the above model expression. Parameter values were drawn uniformly from half to double their median posterior values. While it is often recommended to use quasi-random number samples in calculating these metrics to allow for fewer iterations, our model is sufficiently simple that a uniform distribution can provide an accurate representation of the sensitivity indices. We used 1,000,000 iterations in calculating

S_i	Parameter	S_{T_i}	Parameter
0.4048	d_c	0.6886	d_c
0.0672	p_a	0.1931	d_a
0.0655	d_a	0.1930	p_a
0.0333	p_{hc}	0.1279	p_{hc}
0.0332	λ_c	0.1276	λ_c
0.0324	p_{hh}	0.1268	p_{hh}
0.0035	D	0.0179	D
0.0011	p_f	0.0095	p_f
0.0011	p_{sgf}	0.0093	p_{sgf}

Table 3. Sensitivity analysis of parameters. The first-order sensitivity index and total effect index is given for a sensitivity analysis of nine parameters, with random value variations drawn over 1,000,000 iterations. The output function considered is the probability of illness with salmonellosis due to sheep meat.

these values.

The two indices strongly agree with one another in their ordering of parameter importance, suggesting that most of the impact on the final output by model inputs is through first-order interactions. The parameter d_c , the abattoir processing factor, stands out as the most sensitive parameter. This suggests that in efforts to reduce the microbiological risk load of sheep meat, the most attention and resources should be given to the abattoir module, ensuring that good hygiene is observed, and every effort is made to prevent cross-contamination between carcasses.

Note that the results of the sensitivity analysis show parameters grouped closely together. Unsurprisingly, λ_c , p_{hh} and p_{hc} have incredibly similar sensitivity indices, as these three parameters impact the model in exactly the same way. All are multiplied together in the calculation of λ_e , meaning that an increase to one of the three is no different to increasing either of the other two by the same factor. The same can be said of p_{sgf} and p_f .

Alongside this mathematical definition of sensitivity, we also convey the above sensitivity result by running the model multiple times with different alterations to the initial priors of d_c and m_{sgf} to demonstrate the result shown in Table 3. Fig. 7 shows the resulting posterior distributions for d_c and p_{is} when the Bayesian model is run for varying prior distributions of d_c . Fig. 8 meanwhile shows the resulting posterior distributions for m_{sgf} , p_{sgf} , and p_{is} for different initial prior distributions set on m_{sgf} .

As can be seen in Fig. 7, the posterior of d_c , when converging to a higher value than that shown initially, has a visible impact on the final posterior distribution of p_{is} . However, when we similarly alter the prior distribution around m_{sgf} , as shown in Fig. 8, while we see a significant impact on the distribution of the probability p_{sgf} , this then has no noticeable impact on the final distribution of p_{is} .

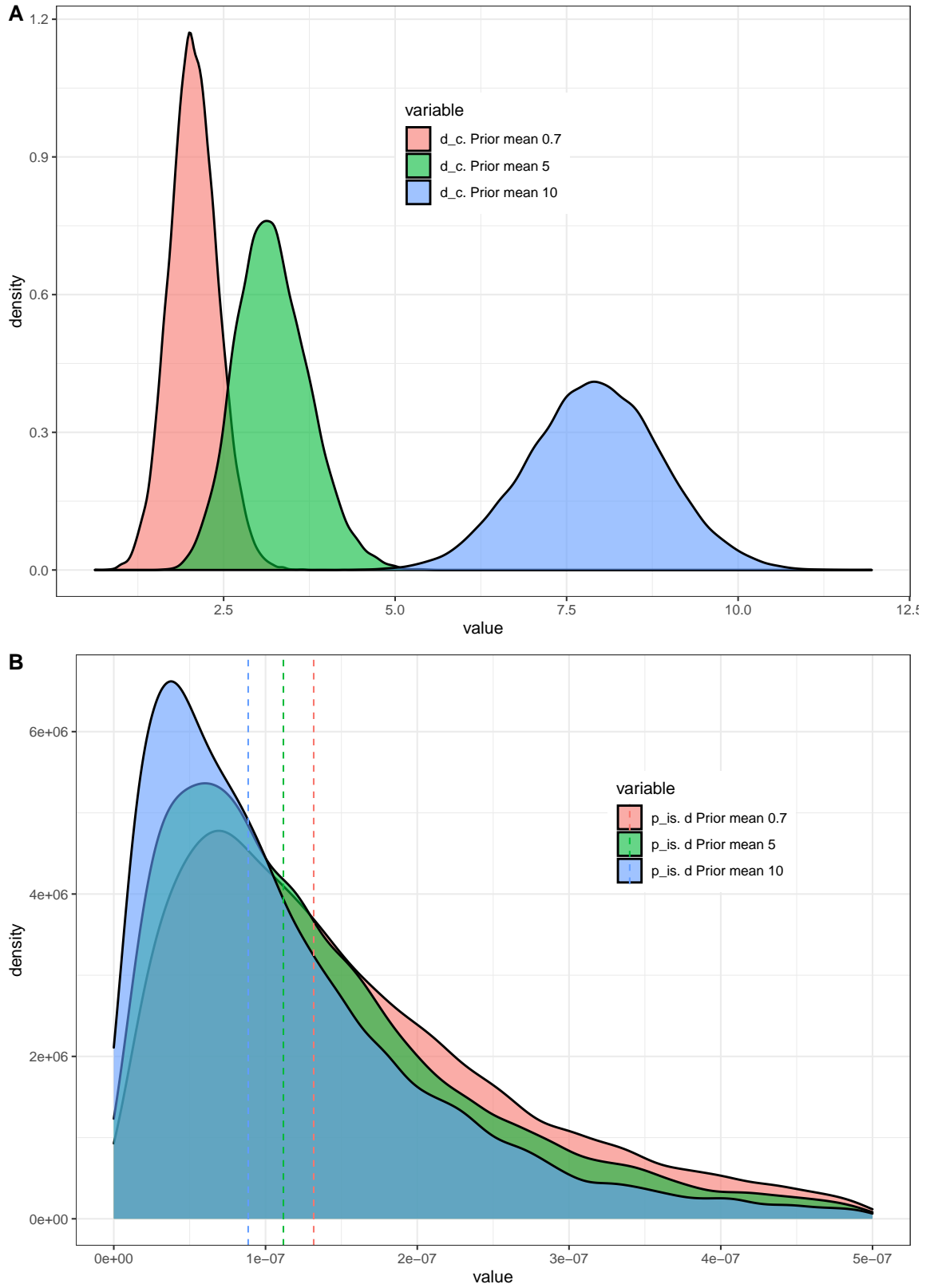


Fig. 7. Sensitivity analysis on the impact of altering the prior set on d_c . The above two plots show the posterior distributions calculated for (A) d_c (B) p_{is} , for different prior distributions set on d_c . The red plots were initialised with $d_c \sim N(0.7, 1.5)$. Green plots with $d_c \sim N(5, 1.5)$. Blue plots with $d_c \sim N(10, 1.5)$. The dashed lines indicate the median values of the associated plots.

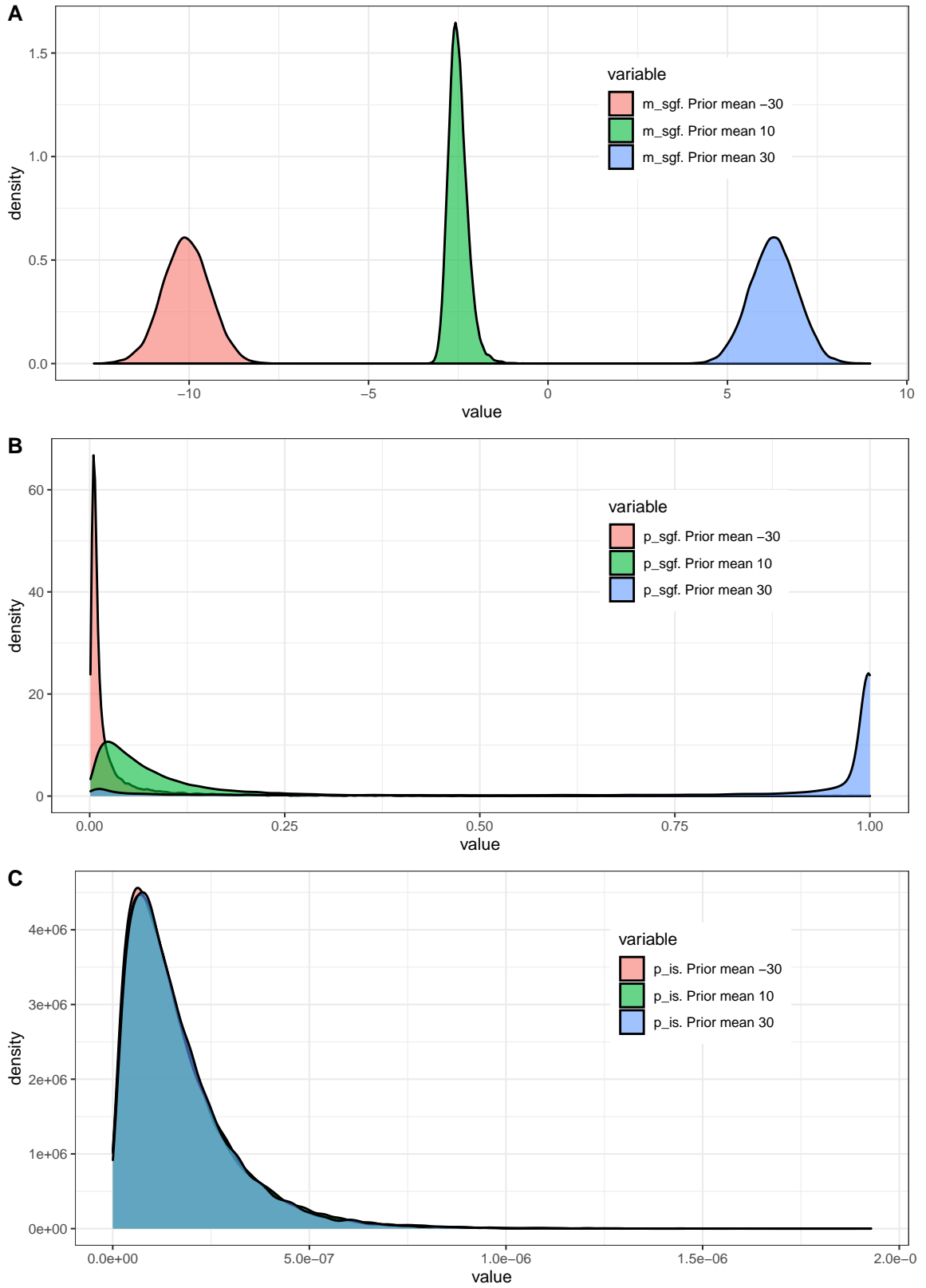


Fig. 8. Sensitivity analysis on the impact of altering the prior set on m_{sgf} . The above three plots show the posterior distributions calculated for (A) m_{sgf} (B) p_{sgf} (C) p_{is} , for different prior distributions set on m_{sgf} . The red plots were initialised with $m_{sgf} \sim N(-30, 1.5)$. Green plots with $m_{sgf} \sim N(10, 1.5)$. Blue plots with $m_{sgf} \sim N(30, 1.5)$.

4 DISCUSSION

Within this project a Bayesian modelling framework has been implemented to build the first (to my knowledge) quantitative microbiological risk assessment model for *Salmonella* in the sheep meat food chain. Data was exhaustively gathered from across the existing literature to inform the resulting posteriors which strongly converged to a final distribution. The final resulting posterior for p_{is} , the probability of becoming ill with salmonellosis via sheep meat in a year, had a median of 1.318×10^{-7} . Assuming a UK population of 66 million, this translates to roughly 8.765 cases of salmonellosis due to sheep meat every year, a feasible result given how 1.45 million cases of foodborne illness in the UK remain unattributed every year (Daniel et al., 2020). The model also capably captured the total number of salmonellosis cases, with the median p_{it} value of 1.669×10^{-4} corresponding to 16.69 salmonellosis cases per 100,000 individuals. Previous PHE records reported an average 15.8 infections per 100,000 individuals in 2007-2016 (Public Health England (PHE), 2018), demonstrating suitable model convergence.

The expanded modules will prove helpful to investigators in demonstrating how robust the sheep meat food chain is to variation in pathogenic load at different stages in the food chain. For example, the sensitivity analysis above revealed that having a greater number of colonised sheep within a flock (given any amount of prior colonisation) (p_{sgf}) is unlikely to account for a significant increase in human illness (Fig. 8). Meanwhile, variation in the abattoir processing factor (d_c) is seen to result in a higher predicted amount of human illness (Fig. 7). The model therefore suggests that, hypothetically, factors in the abattoir processing the colonised meat may play a significant role in future outbreaks, though not to the exclusion of other factors.

Beyond highlighting areas of greatest importance, this work has also brought focus to areas in which further research is required. Greater predictive detail can be provided by modelling the average number of sheep meals, λ_c , hierarchically. Consumer spending habits on sheep meat is likely to be a broad distribution, capturing personal preference and dietary requirements. The NDNS data on population-wide food diaries was of a considerable size ($n = 9,425$); however, due to the short duration of the NDNS data, (only a consecutive four-day period was reported for each individual) it was difficult to accurately observe the spread of lamb consumption across individuals. A high majority of all individuals in the study reported consuming no sheep meat across the four-day period. As such, a hierarchical model formulation would assign an associated high preference for a λ_c value close to 0. In short, the best fit to data of multiple zeros, is a high probability of zero. Taking the average across the whole data revealed a population average of 0.2 sheep meals per four-days. Due to the length of this experiment, only people who ate significantly more than the population average were sufficiently portrayed. As such, any uncertainty in λ_c is introduced only by the resulting posterior distribution, and it is recommended that a longer study period is considered in future experimentation to accurately describe the range of consumer habits.

432 Currently, any foodstuff that is consumed less than 0.25 times per day on average across the population
433 will not be sufficiently reported.

434

435 Similarly, because no data was available to directly inform the cooking module (p_{hc} and p_{hh}), a hi-
436 erarchical structure could not be considered, and uncertainty in these probabilities is only observed in
437 the resulting wide posterior distributions (informed by downstream data). The simplification of this mod-
438 ule to two parameters has scope to be vastly improved by the inclusion of specific cooking practices and
439 thermal inactivation, as considered for other vectors such as pork meat (Swart et al., 2016). Currently
440 no such data unique to sheep meat is available, and it was felt that consumer cooking processes for sheep
441 meat are likely to differ greatly to other meat products. Neves et al. (2018) find that such simplifications
442 of the consumer phase can often result in an overestimation of the impact of consumer interventions,
443 suggesting that the posterior distributions of p_{hc} and p_{hh} may be somewhat overestimates.

444

445 Through compiling data for this model, it became apparent that there was a significant lack of data
446 on microbial prevalence in sheep meat products at the point-of-sale. An accurate calculation of the
447 population attributable risk posed by sheep meat would represent a considerable advancement in the
448 understanding of the sheep meat food chain, and would prove useful for future modelling efforts. The
449 same can be said of better understanding of consumer hygiene practice, with the majority of the current
450 studies into poor food management practice revolving around restricted surveys (Yang et al., 1998). A
451 more precise study into habits within the domestic kitchen would prove useful for food safety work re-
452 gardless of the pathogen of interest.

453

454 Several areas of the model would benefit from expansion as more data becomes available through further
455 investigative study. For example, greater investigation into the beef meat food chain means that greater
456 detail is possible in related QMRAs within cattle. The review of QMRAs describing *Escherichia coli*
457 O157:H7 in beef by Duffy et al. (2006) highlights multiple considered parameters for which descriptive
458 information is simply not yet available relating to sheep meat. Factors such as pathogen concentration
459 in faeces on-farm, temperature abuse, and retail storage times are considered for beef, and the modular
460 nature of our model ensures that such parameters may be incorporated at such a time as they become
461 available.

462

463 Note from the model structure that the probability of *Salmonella* prevalence at retail is informed by
464 observational studies at retail, and modelled as being dependent on prevalence at the end of the slaugh-
465 tering process via the blackbox parameter d_c . While prevalence at these two stages is deeply linked,
466 the model does not explicitly consider the impact of variability of time-temperature profiles, the choice
467 and distribution of meat cuts or specific, separate, decontamination procedures, all of which have been
468 previously identified to impact the concentration of *Salmonella* at retail (Velugoti et al., 2011; Phillips

et al., 2012; Bacon et al., 2000). Our hierarchical model structure ensures that uncertainty and variability amongst these factors is attributed for, however the breadth of these distributions is yet to be epidemiologically validated.

Furthermore, even repeated studies of the parameters currently available, such as rates of colonisation at retail, would allow expansion of the number of considered parameters within the model. For example, the “METZOON” model of *Salmonella* in the pork food chain by Bollaerts et al. (2009) considers 30 distinct parameters within its respective “exposure” module, detailing the hygiene and cooking habits of consumers, compared to the two parameters (p_{hh} and p_{he}) we consider. For our model, with less than a hundred data points in total throughout the model to fit to, such a parameterisation would be entirely incapable of converging to a meaningful conclusion. Hence there is a great need for any and all continued research into *Salmonella* prevalence throughout the sheep meat food chain, through which the presented model can continue to be expanded.

Ideally the results presented above would also be further validated against real epidemiological data, to test the validity of model assumptions and outcomes. Unfortunately, this is not yet possible due to the scarcity of sheep meat outbreak data. The natural progression of this work will be to next assess the findings within against emerging *Salmonella* outbreaks linked to sheep meat. The modular nature of QMRAs means they can easily be further evaluated and expanded as necessary (Heidinger, 2009; Popov et al., 2011).

These caveats notwithstanding, the model has proved elucidating in conveying the burden of *Salmonella* within the sheep meat food chain. Furthermore, the Bayesian formulation ensures that the final predictive results are presented alongside their innate uncertainty, as conveyed through the associated highest density intervals of the resulting posterior distributions. It is hoped that this initial work will prove useful in informing future modelling efforts and experimental studies alike in the ongoing effort to minimise the risk posed to consumers by foodborne pathogens.

CONFLICT OF INTEREST STATEMENT

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

The work was supported through an Engineering and Physical Sciences Research Council (EPSRC) (<https://epsrc.ukri.org/>) Systems Biology studentship award (EP/G03706X/1). The funders had no role

502 in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

503 **ACKNOWLEDGMENTS**

504 This work was conducted as part of a UKRI policy internship at the Food Standards Agency (FSA).

505

506 Thanks are given to Paul Cook, Erin Lewis, and Iulia Gherman of the FSA for discussion and sup-
507 port during the project. Thanks are given to Public Health England and the Animal and Plant Health
508 Agency, whose teams investigating the outbreak of *S. Typhimurium* 3225 helped inform and motivate
509 the work. Thanks are also given to Rob Paton of the Mathematical Ecology Research Group in Oxford
510 for his insight and expertise in model formulation.

References

- I. Albert, E. Grenier, J.-B. Denis, and J. Rousseau. Quantitative risk assessment from farm to fork and beyond: A global Bayesian approach concerning food-borne diseases. *Risk Analysis: An International Journal*, 28(2):557–571, 2008.
- O. Alvseike and E. Skjerve. Prevalence of a *Salmonella* subspecies *diarizonae* in Norwegian sheep herds. *Preventive veterinary medicine*, 52(3-4):277–285, 2002.
- H. Arguello, A. Alvarez-Ordóñez, A. Carvajal, P. Rubio, and M. Prieto. Role of slaughtering in *Salmonella* spreading and control in pork production. *Journal of food protection*, 76(5):899–911, 2013.
- R. Bacon, K. Belk, J. Sofos, R. Clayton, J. Reagan, and G. Smith. Microbial populations on animal hides and beef carcasses at different stages of slaughter in plants employing multiple-sequential interventions for decontamination. *Journal of food protection*, 63(8):1080–1086, 2000.
- D. Beaudequin, F. Harden, A. Roiko, H. Stratton, C. Lemckert, and K. Mengersen. Beyond QMRA: modelling microbial health risk as a complex system using Bayesian networks. *Environment international*, 80:8–18, 2015.
- K. E. Bollaerts, W. Messens, L. Delhalle, M. Aerts, Y. Van der Stede, J. Dewulf, S. Quoilin, D. Maes, K. Mintiens, and K. Grijspeerdt. Development of a quantitative microbial risk assessment for human salmonellosis through household consumption of fresh minced pork meat in Belgium. *Risk Analysis: An International Journal*, 29(6):820–840, 2009.
- R. Bonke, S. Wacheck, C. Bumann, C. Thum, E. Stüber, M. König, R. Stephan, and M. Fredriksson-Ahomaa. High prevalence of *Salmonella enterica* subsp. *diarizonae* in tonsils of sheep at slaughter. *Food Research International*, 45(2):880–884, 2012.
- L. Busani, A. Cigliano, E. Taioli, V. Caligiuri, L. Chiavacci, C. Di Bella, A. Battisti, A. Duranti, M. Gianfranceschi, and M. Nardella. Prevalence of *Salmonella enterica* and *Listeria monocytogenes* contamination in foods of animal origin in Italy. *Journal of food protection*, 68(8):1729–1733, 2005.
- A. Carson and R. Davies. Salmonellosis in sheep. *Veterinary Record*, 183(17):539–539, 2018.
- B. Christensen, H. Sommer, H. Rosenquist, and N. Nielsen. Risk assessment on *Campylobacter jejuni* in chicken products. *The Danish veterinary and food administration, Institute of Food Safety and Toxicology, Division of Microbiological safety.*, 2001.
- N. Daniel, N. Casadevall, P. Sun, D. Sugden, and V. Aldin. The burden of foodborne disease in the UK 2018. *London: FSA*, 2020.
- R. Davies, R. Dalziel, J. Gibbens, J. Wilesmith, J. Ryan, S. Evans, C. Byrne, G. Paiba, S. Pascoe, and C. Teale. National survey for *Salmonella* in pigs, cattle and sheep at slaughter in Great Britain (1999–2000). *Journal of Applied Microbiology*, 96(4):750–760, 2004.

544 Department for Environment, Food & Rural Affairs (DEFRA). Literature review & survey of conditions
545 relevant to farm animal welfare in lairages - MH0132. Available at [http://sciencesearch.defra.](http://sciencesearch.defra.gov.uk/)
546 [gov.uk/](http://sciencesearch.defra.gov.uk/) (Accessed on 31 October 2019), 2006.

547 R. M. Dorazio. Bayesian data analysis in population ecology: motivations, methods, and benefits. *Pop-*
548 *ulation ecology*, 58(1):31–44, 2016.

549 E. Duffy, K. Belk, J. Sofos, S. LeValley, M. Kain, J. Tatum, G. Smith, and C. Kimberling. Microbial
550 contamination occurring on lamb carcasses processed in the United States. *Journal of Food Protection*,
551 64(4):503–508, 2001.

552 G. Duffy, E. Cummins, P. Nally, S. O’Brien, and F. Butler. A review of quantitative microbial risk
553 assessment in the management of escherichia coli o157: H7 on beef. *Meat science*, 74(1):76–88, 2006.

554 L. Duffy, A. Small, and N. Fegan. Concentration and prevalence of Escherichia coli O157 and Salmonella
555 serotypes in sheep during slaughter at two Australian abattoirs. *Australian veterinary journal*, 88(10):
556 399–404, 2010.

557 B. Efron. Bayesians, frequentists, and scientists. *Journal of the American Statistical Association*, 100
558 (469):1–5, 2005.

559 H. Esmaili and H. K. Rahmani. Detection of Salmonella carriers in sheep and goat flocks of Bushehr
560 and Lorestan provinces, Iran. *Journal of Medical Bacteriology*, pages 50–53, 2016.

561 European Food Safety Authority (EFSA). A quantitative microbiological risk assessment on Salmonella
562 in meat: Source attribution for human salmonellosis from meat. *The EFSA Journal (2008)*, 625:1–32,
563 2008.

564 M. Evans, R. Salmon, L. Nehaul, S. Mably, L. Wafford, M. Nolan-Farrell, D. Gardner, and C. Ribeiro.
565 An outbreak of Salmonella typhimurium DT170 associated with kebab meat and yoghurt relish. *Epi-*
566 *demiology & Infection*, 122(3):377–383, 1999.

567 Food Standards Agency (FSA). Project B18018: A UK-wide survey of microbiological contamination of
568 fresh red meats on retail sale, 2010.

569 A. Gelman. Prior distributions for variance parameters in hierarchical models (comment on article by
570 Browne and Draper). *Bayesian analysis*, 1(3):515–534, 2006.

571 F. Grau and M. Smith. Salmonella contamination of sheep and mutton carcasses related to pre-slaughter
572 holding conditions. *Journal of Applied Bacteriology*, 37(1):111–116, 1974.

573 C. N. Haas. Estimation of risk due to low doses of microorganisms: a comparison of alternative method-
574 ologies. *American journal of epidemiology*, 118(4):573–582, 1983.

575 K. E. Hanlon, M. F. Miller, L. M. Guillen, A. Echeverry, E. Dormedy, B. Cemo, L. A. Branham,
576 S. Sanders, and M. M. Brashears. Presence of Salmonella and Escherichia coli O157 on the hide,
577 and presence of Salmonella, Escherichia coli O157 and Campylobacter in feces from small-ruminant
578 (goat and lamb) samples collected in the United States, Bahamas and Mexico. *Meat science*, 135:1–5,
579 2018.

580 J. C. Heidinger. *Quantitative microbial risk assessment (QMRA) development: QMRA for Staphylococcus*
581 *aureus and Staphylococcus enterotoxin A in raw milk as an example*. University of California, Davis,
582 2009.

583 S. Hjartardóttir, E. Gunnarsson, and J. Sigvaldadóttir. Salmonella in sheep in Iceland. *Acta Veterinaria*
584 *Scandinavica*, 43(1):43, 2002.

585 D. Kane. The prevalence of Salmonella infection in sheep at slaughter. *New Zealand veterinary journal*,
586 27(6):110–113, 1979.

587 D. Kilsby and M. Pugh. The relevance of the distribution of micro-organisms within batches of food
588 to the control of microbiological hazards from foods. *Journal of Applied Bacteriology*, 51(2):345–354,
589 1981.

590 J. Kruschke. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press, 2014.

591 S. Kumar, S. Saxena, and B. Gupta. Carrier rate of Salmonellas in sheep and goats and its public health
592 significance. *Epidemiology & Infection*, 71(1):43–48, 1973.

593 C. Little, J. Richardson, R. Owen, E. De Pinna, and E. Threlfall. Campylobacter and Salmonella in raw
594 red meats in the United Kingdom: prevalence, characterization and antimicrobial resistance pattern,
595 2003–2005. *Food microbiology*, 25(3):538–543, 2008.

596 U. Methner and U. Moog. Occurrence and characterisation of Salmonella enterica subspecies diarizonae
597 serovar 61: k: 1, 5,(7) in sheep in the federal state of Thuringia, Germany. *BMC veterinary research*,
598 14(1):401, 2018.

599 A. Milnes, I. Stewart, F. Clifton-Hadley, R. Davies, D. Newell, A. Sayers, T. Cheasty, C. Cassar, A. Ridley,
600 A. Cook, et al. Intestinal carriage of verocytotoxigenic Escherichia coli O157, Salmonella, thermophilic
601 Campylobacter and Yersinia enterocolitica, in cattle, sheep and pigs at slaughter in Great Britain
602 during 2003. *Epidemiology & Infection*, 136(6):739–751, 2008.

603 M. I. Neves, S. N. Mungai, and M. J. Nauta. Can stochastic consumer phase models in qmra be simplified
604 to a single factor? *Microbial Risk Analysis*, 8:53–60, 2018.

605 P. Nottingham and A. Urselmann. Salmonella infection in calves and other animals. *New Zealand journal*
606 *of agricultural research*, 4(5-6):449–460, 1961.

607 Organisation for Economic Cooperation and Development (OECD). Meat consumption (indicator). doi:
608 10.1787/fa290fd0-en (Accessed on 17 October 2019), 2019.

609 D. Phillips, J. Sumner, J. F. Alexander, and K. M. Dutton. Microbiological quality of Australian sheep
610 meat. *Journal of food protection*, 64(5):697–700, 2001.

611 D. Phillips, K. Bridger, I. Jenson, and J. Sumner. An australian national survey of the microbiological
612 quality of frozen boneless beef and beef primal cuts. *Journal of food protection*, 75(10):1862–1866,
613 2012.

614 M. Plummer. Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. Available
615 at: <http://mcmc-jags.sourceforge.net/>, 2007.

616 M. Plummer, A. Stukalov, and M. Denwood. rjags: Bayesian graphical models using MCMC. Available
617 at: <http://CRAN.R-project.org/package=rjags>. R package version, 2016.

618 V. Popov, H. L. Lauzon, M. Haque, F. Leroi, R. Gospavic, et al. Evaluation of qmra performance for
619 listeria monocytogenes in cold smoked salmon. *WIT Transactions on Biomedicine and Health*, 15:
620 197–297, 2011.

621 Public Health England (PHE). Salmonella data 2007 to 2016. Available at [https://www.gov.](https://www.gov.uk/government/publications/salmonella-national-laboratory-data)
622 [uk/government/publications/salmonella-national-laboratory-data](https://www.gov.uk/government/publications/salmonella-national-laboratory-data) (Accessed on 17 October
623 2019), 2018.

624 G. Purvis, K. Hullah, S. Pascoe, S. Evans, and R. Davies. Persistence of Salmonella Typhimurium DT120
625 in abattoir paddocks holding sheep. *Veterinary record*, 157(6):165–167, 2005.

626 A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola.
627 *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

628 J. Samuel, J. Eccles, and J. Francis. Salmonella in the intestinal tract and associated lymph nodes of
629 sheep and cattle. *Epidemiology & Infection*, 87(2):225–232, 1981.

630 M. Sandberg, O. Alvseike, and E. Skjerve. The prevalence and dynamics of Salmonella enterica IIIb 61:
631 k: 1, 5,(7) in sheep flocks in Norway. *Preventive veterinary medicine*, 52(3-4):267–275, 2002.

632 P. J. Schmidt, K. D. Pintar, A. M. Fazil, and E. Topp. Harnessing the theoretical foundations of the
633 exponential and beta-poisson dose-response models to quantify parameter uncertainty using markov
634 chain monte carlo. *Risk Analysis*, 33(9):1677–1693, 2013.

635 A. Small, C.-A. Reid, S. Avery, N. Karabasil, C. Crowley, and S. Buncic. Potential for the spread of
636 Escherichia coli O157, Salmonella, and Campylobacter in the lairage environment at abattoirs. *Journal*
637 *of food protection*, 65(6):931–936, 2002.

638 W. Smerdon, G. Adak, S. O'Brien, I. Gillespie, and M. Reacher. General outbreaks of infectious intestinal
639 disease linked with red meat, England and Wales, 1992-1999. *Communicable Disease and Public Health*,
640 4(4):259-267, 2001.

641 J. Smid, R. de Jonge, A. H. Havelaar, and A. Pielaat. Variability and uncertainty analysis of the cross-
642 contamination ratios of Salmonella during pork cutting. *Risk Analysis*, 33(6):1100-1115, 2013.

643 K. Sörén, M. Lindblad, C. Jernberg, E. Eriksson, L. Melin, H. Wahlström, and M. Lundh. Changes in the
644 risk management of Salmonella enterica subspecies diarizonae serovar 61:(k): 1, 5,(7) in Swedish sheep
645 herds and sheep meat due to the results of a prevalence study 2012. *Acta Veterinaria Scandinavica*,
646 57(1):6, 2015.

647 A. Swart, F. Van Leusden, and M. Nauta. A qmra model for salmonella in pork products during
648 preparation and consumption. *Risk Analysis*, 36(3):516-530, 2016.

649 M. Synnott, D. Morse, H. Maguire, F. Majid, M. Plummer, M. Leicester, E. Threlfall, and J. Cowden.
650 An outbreak of Salmonella mikawasima associated with doner kebabs. *Epidemiology & Infection*, 111
651 (3):473-482, 1993.

652 The World Health Organization (WHO). *Risk assessments of Salmonella in eggs and broiler chickens*,
653 volume 1. World Health Organization, 2002.

654 P. B. Vanderlinde, B. Shay, and J. Murray. Microbiological status of Australian sheep meat. *Journal of*
655 *Food Protection*, 62(4):380-385, 1999.

656 B. Vanselow, M. Hornitzky, K. Walker, G. Eamens, G. Bailey, P. Gill, K. Coates, B. Corney, J. Cronin,
657 and S. Renilson. Salmonella and on-farm risk factors in healthy slaughter-age cattle and sheep in
658 eastern Australia. *Australian Veterinary Journal*, 85(12):498-502, 2007.

659 P. R. Velugoti, L. K. Bohra, V. K. Juneja, L. Huang, A. L. Wesseling, J. Subbiah, and H. Thippareddi.
660 Dynamic model for predicting growth of salmonella spp. in ground sterile pork. *Food Microbiology*, 28
661 (4):796-803, 2011.

662 F. A. Verdonck, J. Jaworska, O. Thas, and P. A. Vanrolleghem. Determining environmental standards
663 using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study. *Analytica*
664 *Chimica Acta*, 446(1-2):427-436, 2001.

665 T. L. Wong, C. Nicol, R. Cook, and S. MacDiarmid. Salmonella in uncooked retail meats in New Zealand.
666 *Journal of food protection*, 70(6):1360-1365, 2007.

667 R. Yang, C. Jacobson, G. Gardner, I. Carmichael, A. J. Campbell, and U. Ryan. Longitudinal prevalence,
668 faecal shedding and molecular characterisation of Campylobacter spp. and Salmonella enterica in sheep.
669 *The Veterinary Journal*, 202(2):250-254, 2014.

670 S. Yang, M. G. Leff, D. McTague, K. A. Horvath, J. Jackson-Thompson, T. Murayi, G. K. Boeselager,
671 T. A. Melnik, M. C. Gildemaster, and D. L. Ridings. Multistate surveillance for food-handling, prepara-
672 tion, and consumption behaviors associated with foodborne diseases: 1995 and 1996 BRFSS food-safety
673 questions. *Morbidity and Mortality Weekly Report: CDC Surveillance Summaries*, pages 33–57, 1998.

A Appendices

A.1 Appendix 1 - Bayesian Statistics

This brief section aims to convey the basic principles of Bayesian statistics, and familiarise the reader with the terminology that is be used throughout the manuscript. For an in-depth explanation, I recommend the text by Kruschke (2014).

Bayesian statistics is derived wholly from the relationship defined by Bayes' theorem,

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}. \quad (1)$$

If we consider θ as some statistical parameter we wish to infer, and D as some data informing the parameter, then equation (1) expresses that the probability distribution for our value of θ , given our dataset ($P(\theta|D)$), is proportional to the **likelihood** of such data ($P(D|\theta)$) multiplied by the probability distribution of θ free of any data ($P(\theta)$).

Spoken plainly, one starts with a **prior** probabilistic understanding of the values θ , often informed by expert opinion, and by utilising relevant data, D , we update our belief in the values θ may take, producing a new **posterior** distribution. Mnemonically, if we wished to calculate the probability that a flipped coin will land heads up, we may have a **prior** belief that the coin is fair. However, upon observing a data set of 5 coin flips, all of which produced heads, we may update our **posterior** belief to reflect that the coin may be biased.

The analytical difficulty in this calculation lies in computing $P(D) = \int P(D|\theta)P(\theta)d\theta$, which is often near impossible for realistically complex models. Fortunately modern computing power enables us to efficiently estimate our posterior distributions through algorithms such as Gibbs sampling and other Metropolis-Hastings schemes.

Hierarchical systems represent multi-variable models where some parameters depend on other parameters. Returning to the example of a coin flip, say the probability of heads (θ) is dependent on the factory in which the coin was minted. The probability that a coin was from a certain factory (ω) will then inform our value of (θ). Expressed mathematically, equation (1) now becomes:

$$\begin{aligned} P(\theta, \omega|D) &= \frac{P(D|\theta, \omega)P(\theta, \omega)}{P(D)} \\ &= \frac{P(D|\theta, \omega)P(\theta|\omega)P(\omega)}{P(D)}. \end{aligned} \quad (2)$$

702 This means that a prior distribution is only required for ω , as this distribution will directly inform our
703 **conditional prior** of θ , via our model formulation. As such, when provided with data on coin flips from
704 multiple coins from different factories, we obtain a posterior probability distribution of which factory
705 a coin has come from, and the resulting probability of a coin flip resulting in heads. This structure of
706 conditional independence means that data relating specifically to one parameter can still help inform the
707 posterior of all other dependent variables, a key advantage of Bayesian inference.

708

709 Based on this knowledge, hierarchical models are a naturally appropriate way to simulate risk assess-
710 ments, as naturally we are interested in tracing back a series of dependencies. The probability of a
711 sheep being colonised by *Salmonella* on the farm will directly influence the probability of its carcass
712 being colonised after slaughter, which in turn will directly influence the probability of a consumer falling
713 ill. Furthermore, relevant data for understanding the food chain comes from a large number of studies,
714 of varying location and experimental design, focusing on different specific parts of the food chain. A
715 Bayesian approach means that data on one specific module will help inform the parameters across the
716 entire model, not just those in the respective module.

717

718 As a final point, the logit function, $\text{logit}(x) = \log\left(\frac{x}{1-x}\right)$, is employed throughout the model. This
719 formulation is often used in hierarchical models to bound a resulting parameter to be between 0 and 1.
720 Note from the formulation that for any non-unity $x \in \mathbb{R}$, $\text{logit}(x) \in (0, 1]$. This is a particularly useful
721 formulation when dependent variables and factors may not themselves be probabilities, and as such can-
722 not be cleanly expressed in a closed-form expression. For example, if a probability p , was dependent on
723 real-valued factors i, j, k , one could express p as $p = \text{logit}(i + j + k)$.

724 **A.2 Appendix 2 - Binomial count data**

725 This appendix explicitly lists the binomial data taken from the literature used to fit the model via
726 Binomial likelihood distributions.

Related variable	Source	Positive samples	Total samples
On-farm			
p_f	Alvseike and Skjerve (2002)	16	133
p_f	Hjartardóttir et al. (2002)	0/3/5	50/43/39
p_f	Sörén et al. (2015)	57	244
p_f	Sandberg et al. (2002)	7	50
p_f	Yang et al. (2014)	8	8
p_f	Vanselow et al. (2007)	2/4	48/46
p_f	Methner and Moog (2018)	74/20	90/31
p_{sgf}	Hjartardóttir et al. (2002)	4 [†]	6
p_{sgf}	Vanselow et al. (2007)	15 [†]	25
p_{sgf}	Hjartardóttir et al. (2002)	1/1/1/1/2	2/2/2/2/2
p_{sgf}	Sandberg et al. (2002)	1/1/1/1/1/2/3	23/26/28/30/32/28/35
		4/2/2/3/15/4/4	163/163/170/162/157/149/171
p_{sgf}	Yang et al. (2014)*	/4/5/10/9 4/29/38/ 6/5/8/13/5/1/4/2	164/159/178/173/129/159/160 170/127/122/122/107/49/102/103
p_s	Hjartardóttir et al. (2002)	3 [†]	229
p_s	Vanselow et al. (2007)	10 [†]	470
p_s	Hjartardóttir et al. (2002)	0/6	2592/72
p_s	Sandberg et al. (2002)	10	1233
p_s	Esmaeili and Rahmani (2016)	2/0	189/200
Meat processing			
p_a	Grau and Smith (1974)	0/0/0/3/0/0/1 0/0/2/3/3/0/1	10/10/10/10/10/10/10 10/10/10/10/10/10/10
p_{as}	Hanlon et al. (2018)	32/5/0/4/33	78/25/98/45/286
p_{as}	Sörén et al. (2015)	3/2	605/404
p_{as}	Davies et al. (2004)	1/72	973/653
p_{as}	Bonke et al. (2012)	43	100
p_{as}	Nottingham and Urselmann (1961)	5	33
p_{as}	Samuel et al. (1981)	43/30	100/40
p_{as}	Kumar et al. (1973)	25/26	812/683
p_{as}	Duffy et al. (2010)	20/12/0/0/0	40/24/30/34/35
p_c	Kane (1979)	96	2027
p_c	Milnes et al. (2008)	31	2825
p_c	Duffy et al. (2001)	24/15	1262/1260
p_c	Duffy et al. (2010)	1/1/0/0/0	40/24/30/30/35
p_c	Little et al. (2008)	13/5	744/161
p_c	Vanderlinde et al. (1999)	27/27	470/415
p_c	Phillips et al. (2001)	1/6	921/469
p_c	Busani et al. (2005)	0	151
p_c	Wong et al. (2007)	3	230
p_c	Food Standards Agency (FSA) (2010)	0	1056
Illness			
p_{it}	Public Health England (PHE) (2018)	22/19/17/15/15 14/13/13/15/15	100,000 for each

* Back-inferred from mean and 95% CI.

[†] pooled samples, as detailed in manuscript.

Table A1. Exhaustive binomial observation data used to fit the model.

727 A.3 Appendix 3 - Posterior diagnostic plots

728 We include the diagnostic and density plots for all model variables, beyond just those presented in Figs.
729 2 - 6 of the manuscript.

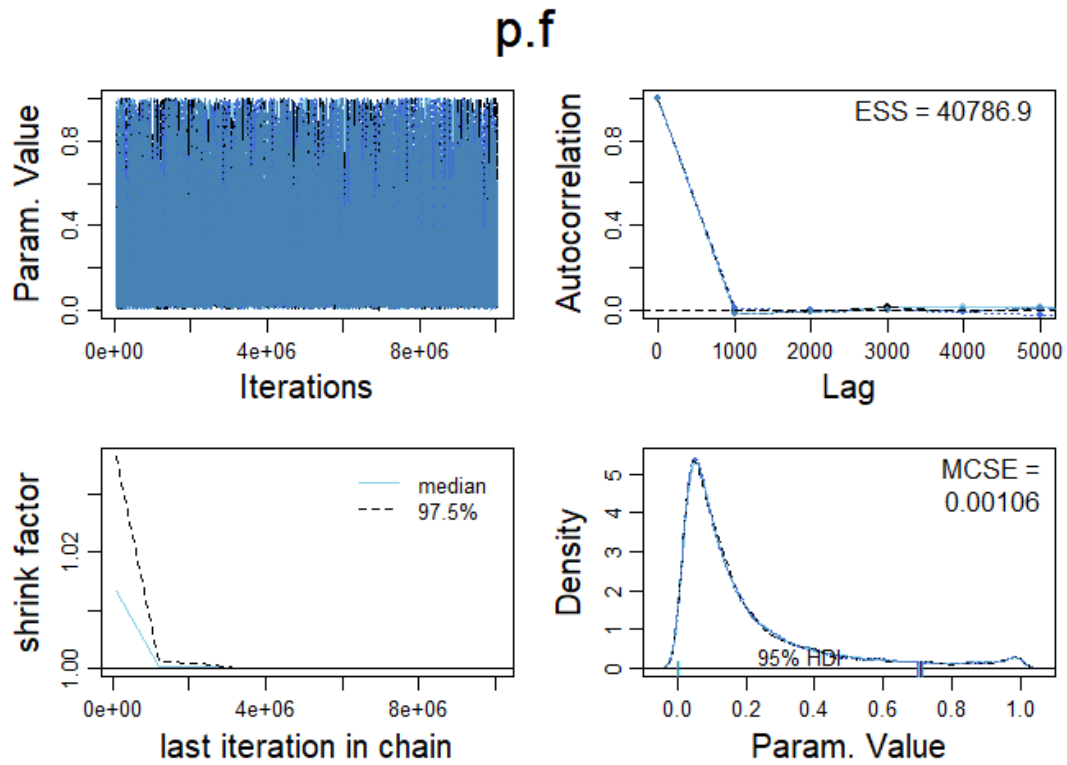


Fig. A1. Posterior of $p.f$. Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

m.f

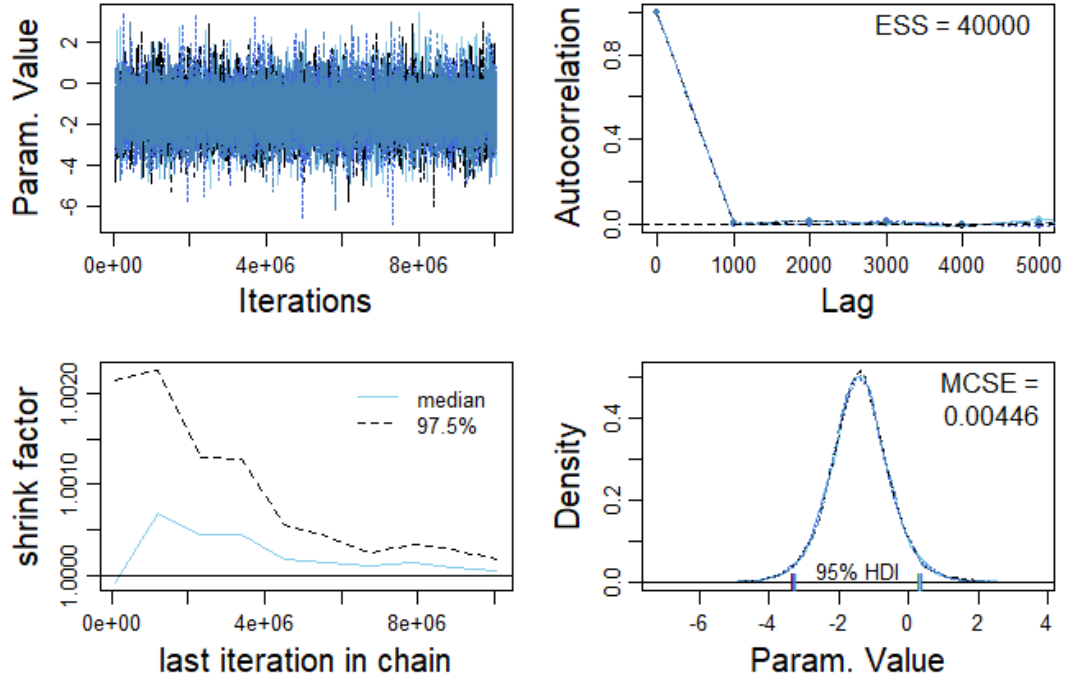


Fig. A2. Posterior of m_f . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

s.f

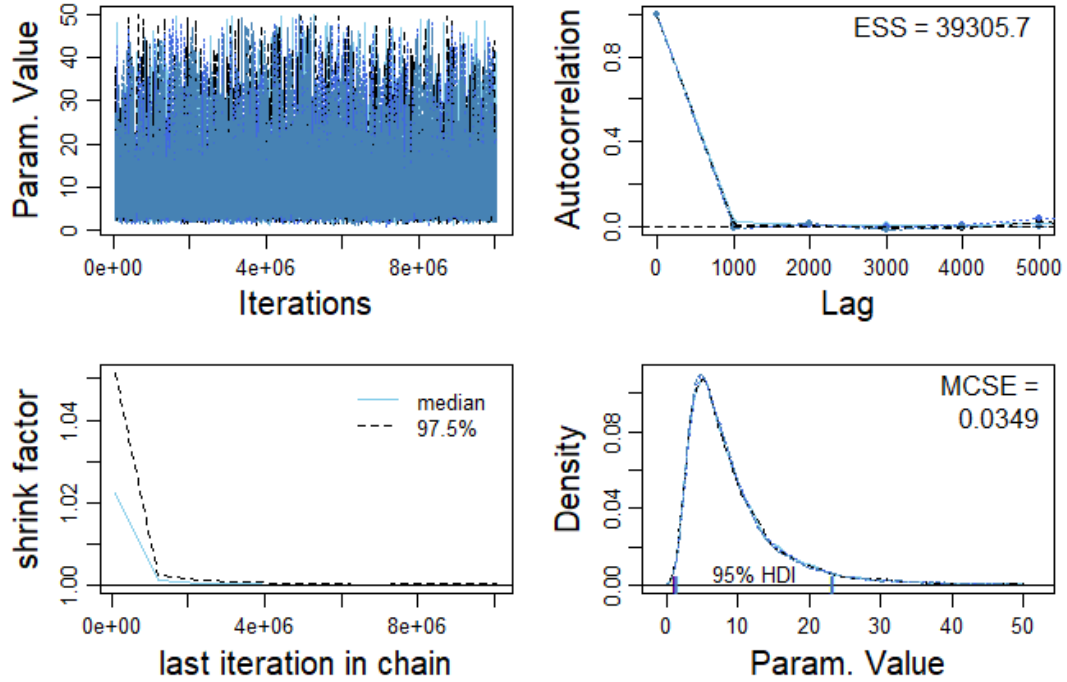


Fig. A3. Posterior of s_f . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.sgf

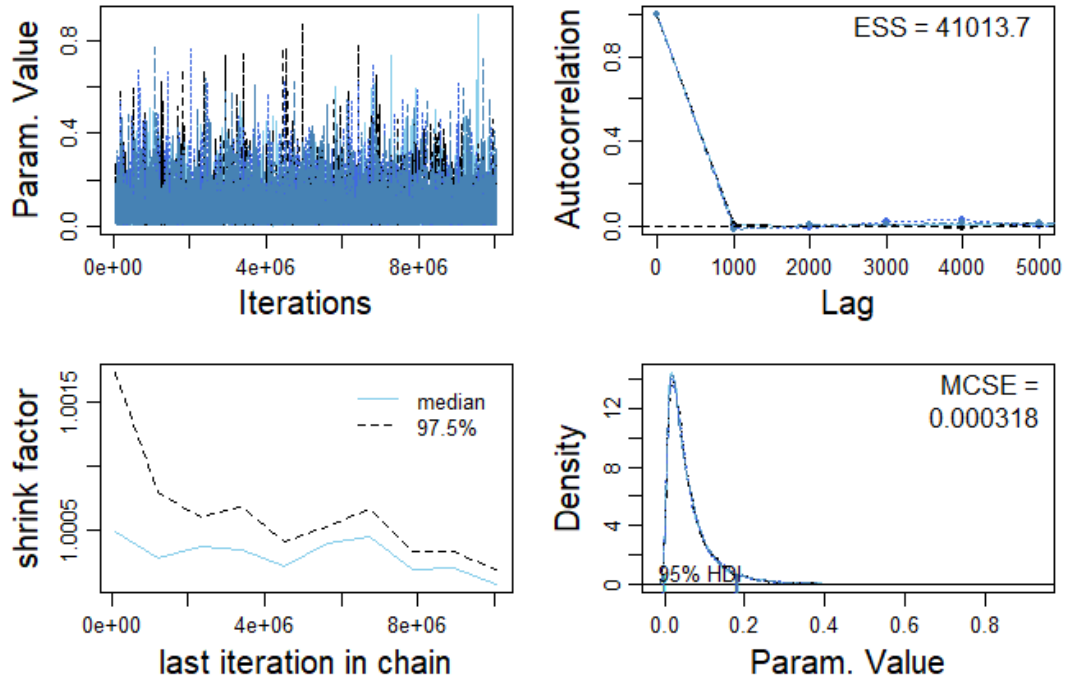


Fig. A4. Posterior of p_{sgf} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

m.sgf

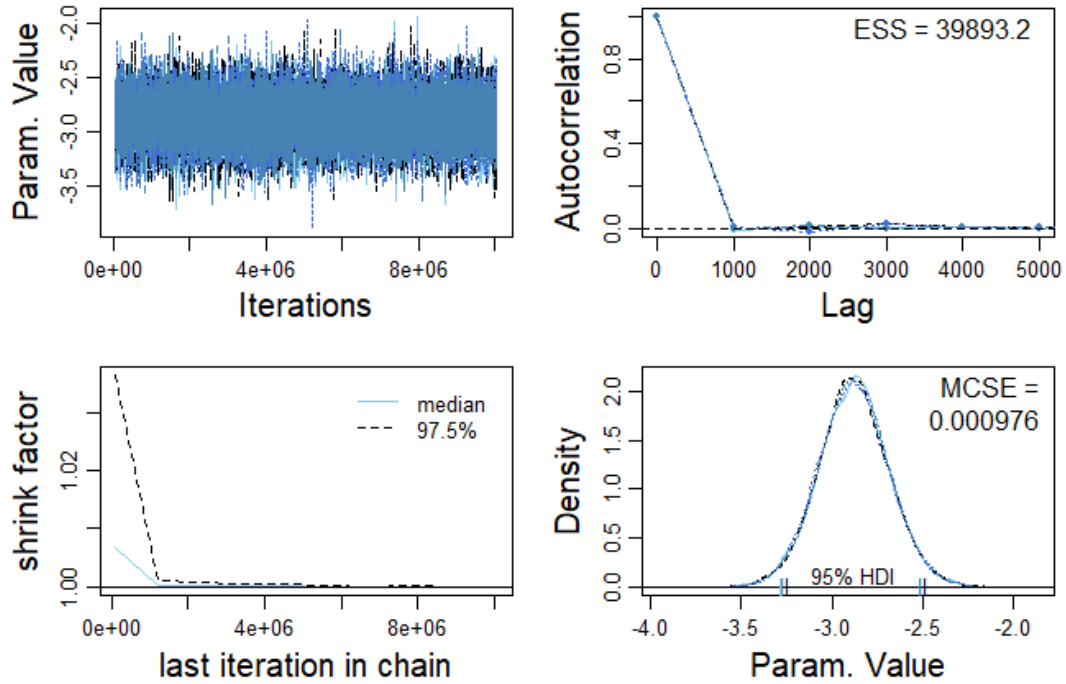


Fig. A5. Posterior of m_{sgf} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

s.sgf

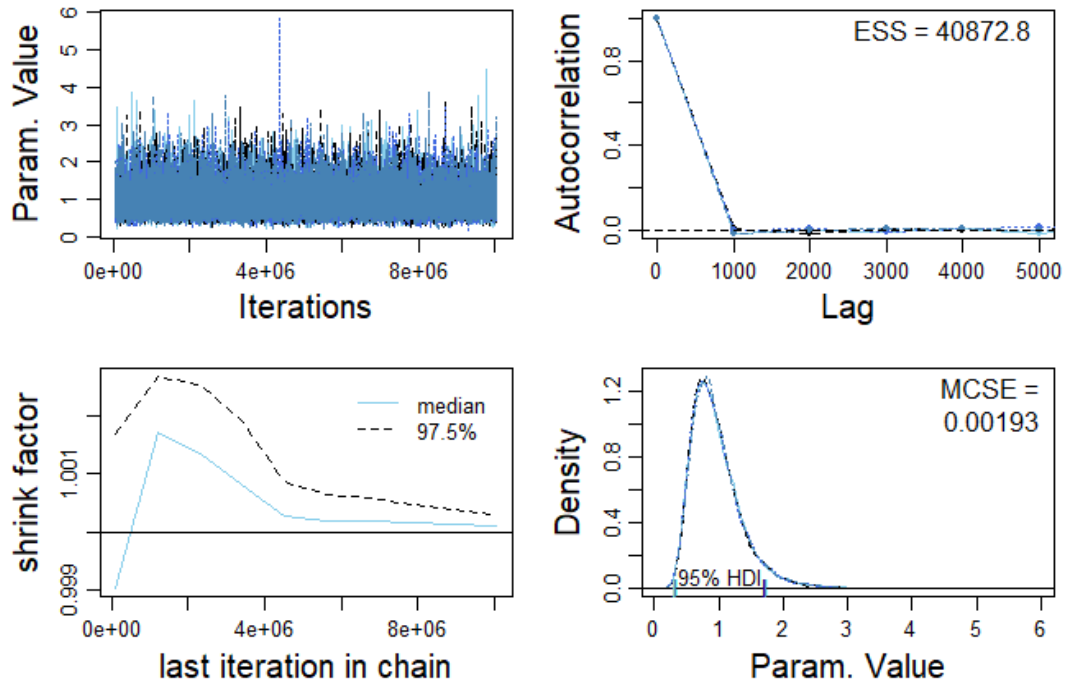


Fig. A6. Posterior of s_{sgf} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.s

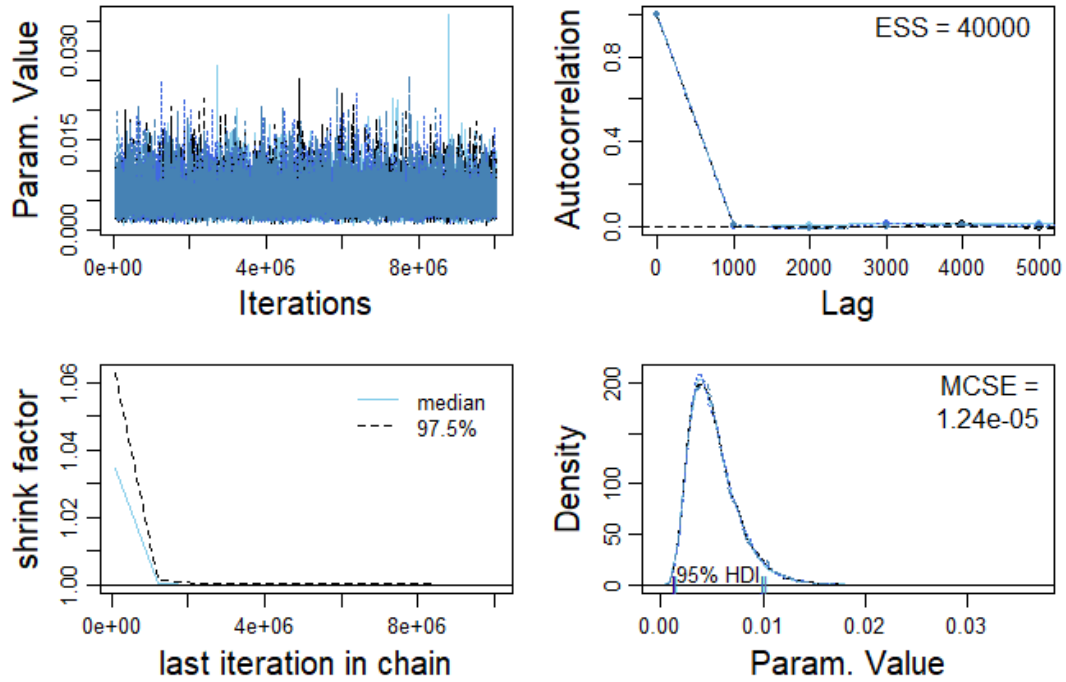


Fig. A7. Posterior of p_s . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.a

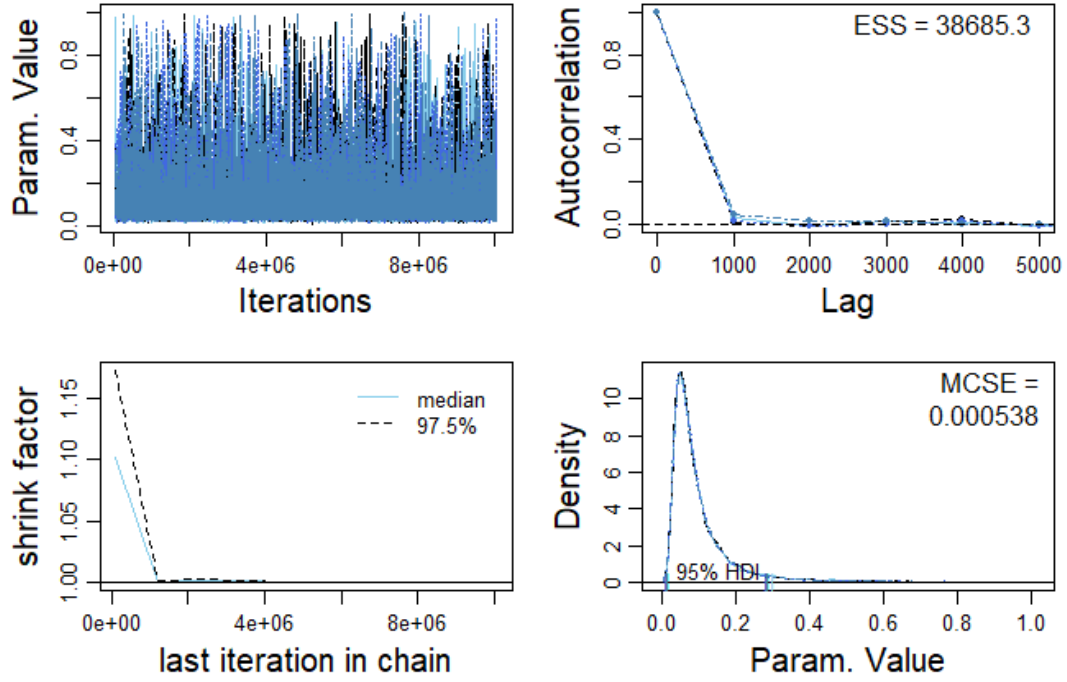


Fig. A8. Posterior of p_a . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

m.a

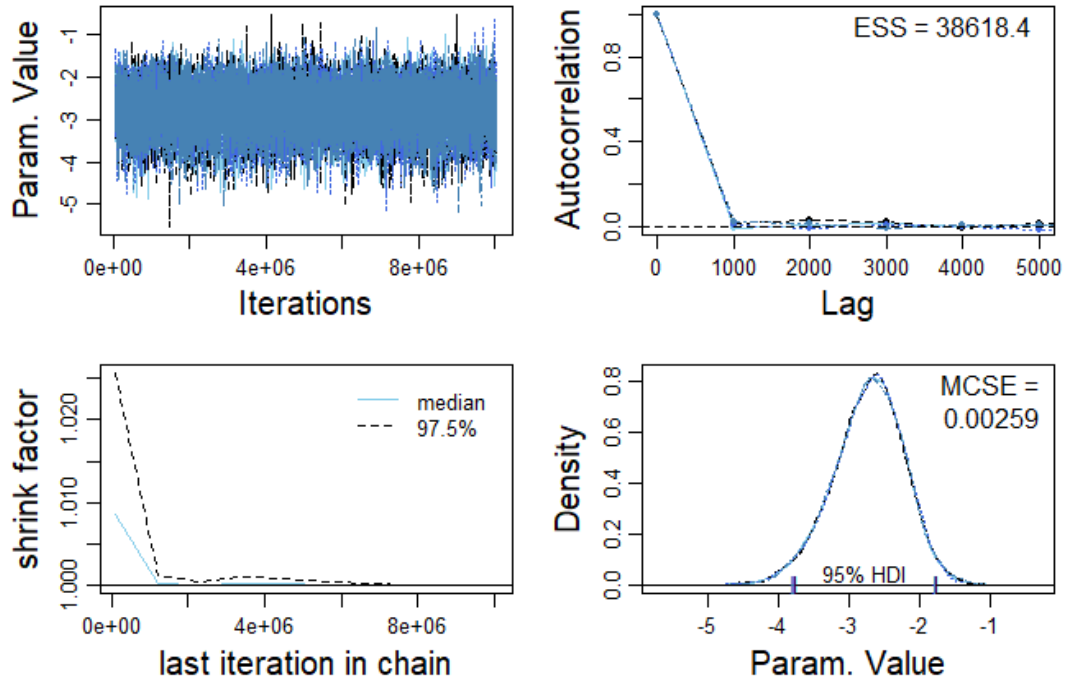


Fig. A9. Posterior of m_a . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

s.a

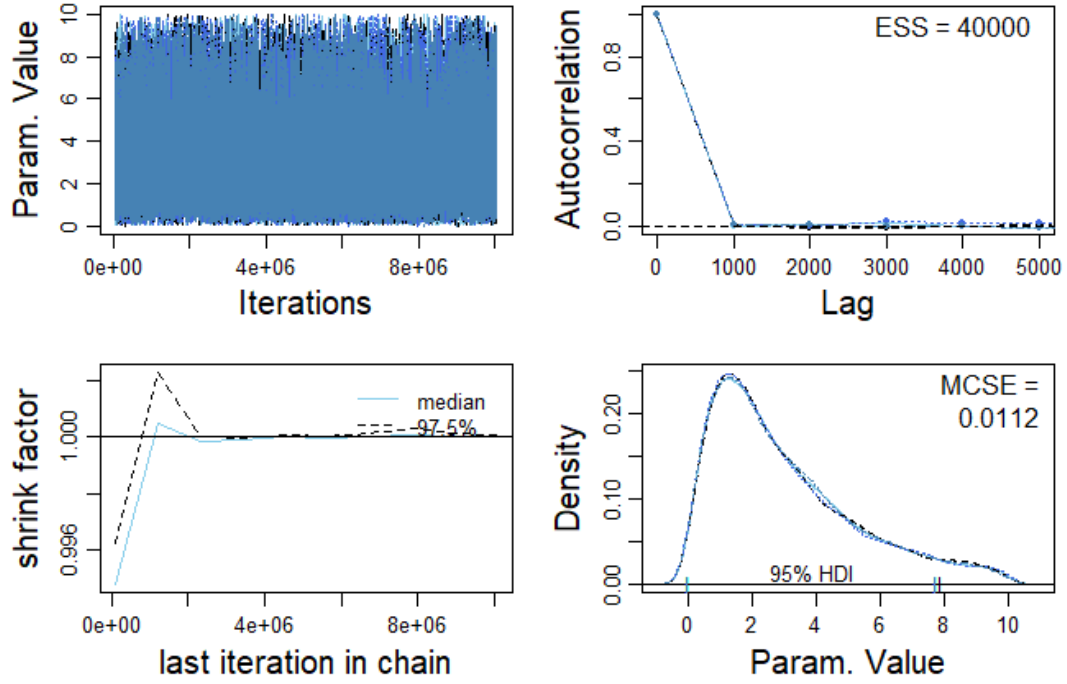


Fig. A10. Posterior of s_a . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.as

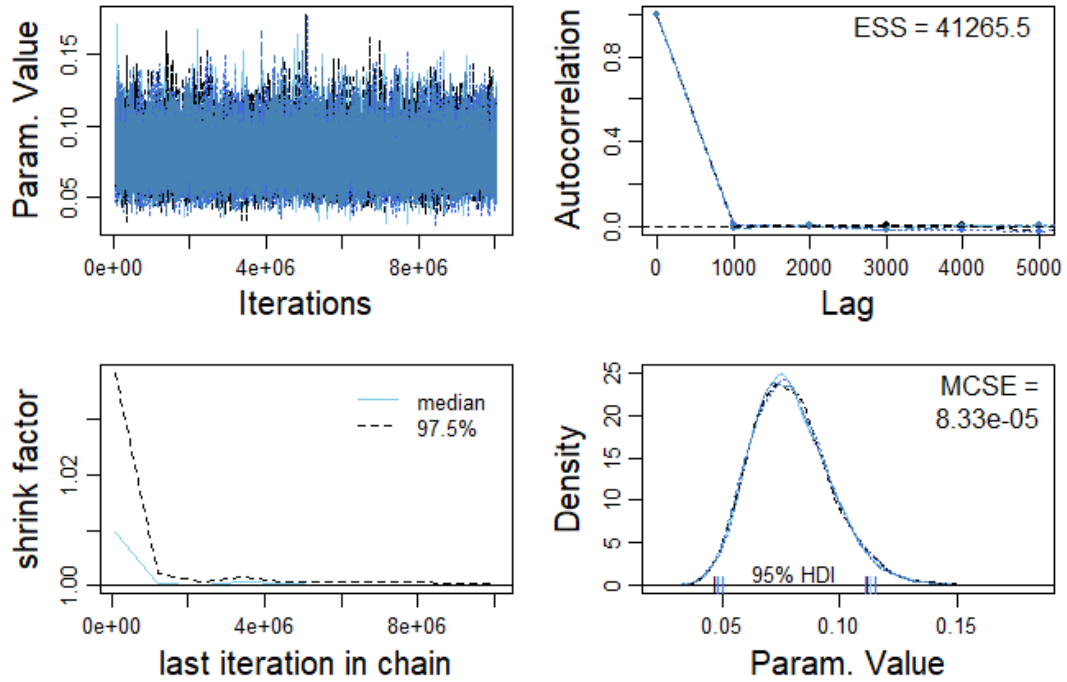


Fig. A11. Posterior of p_{as} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

d.a

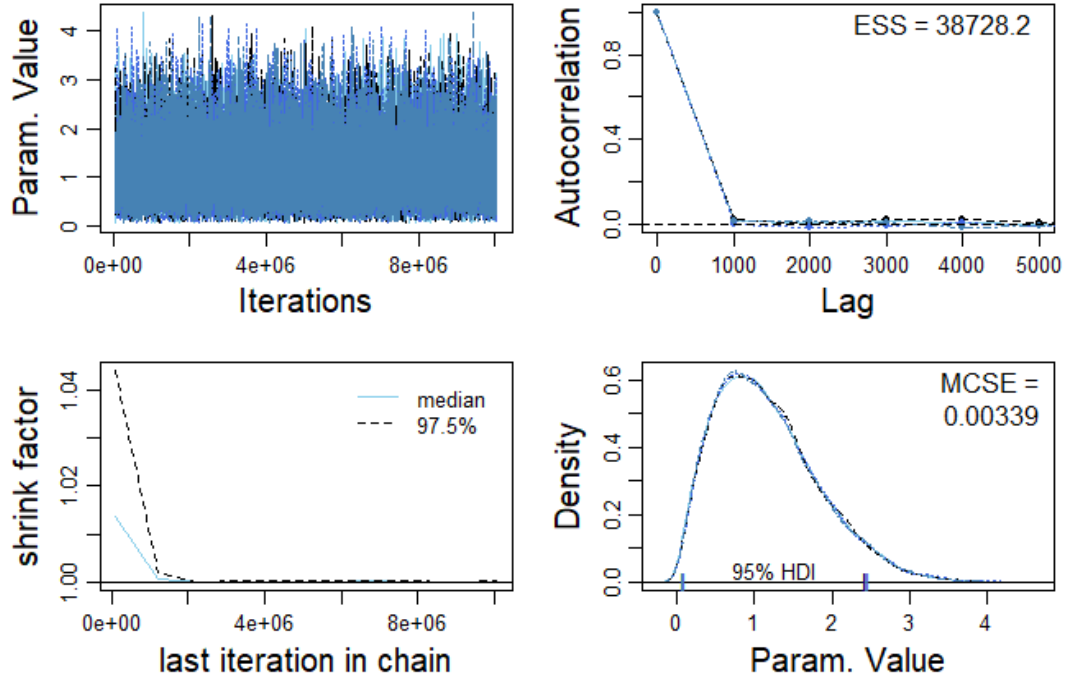


Fig. A12. Posterior of d_a . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.c

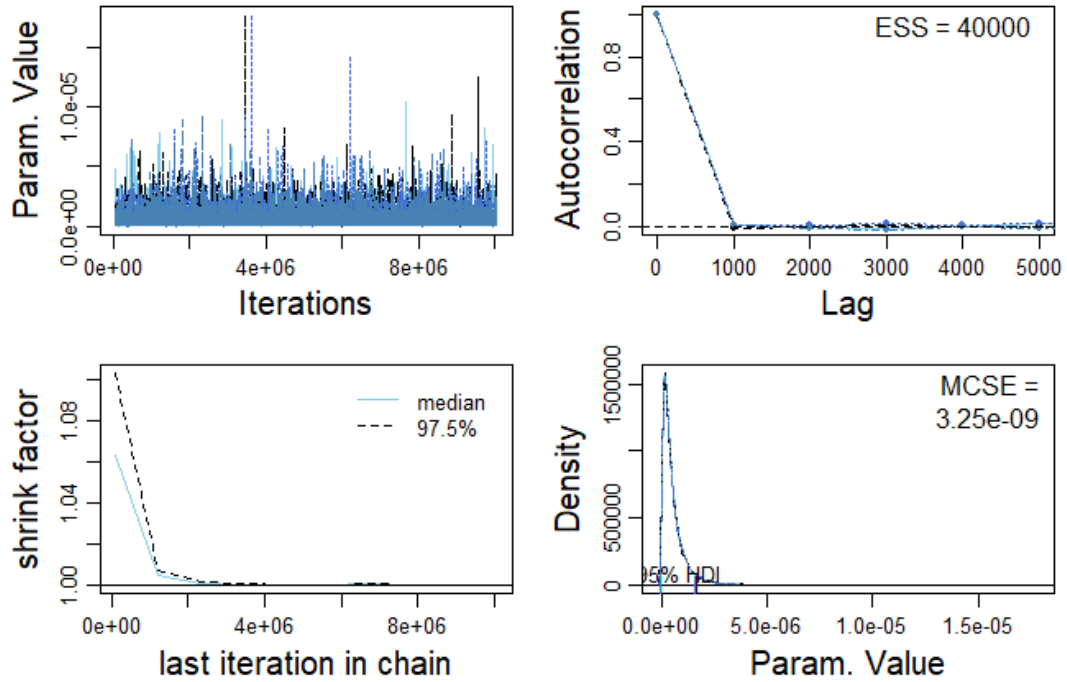


Fig. A13. Posterior of p_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

m.c

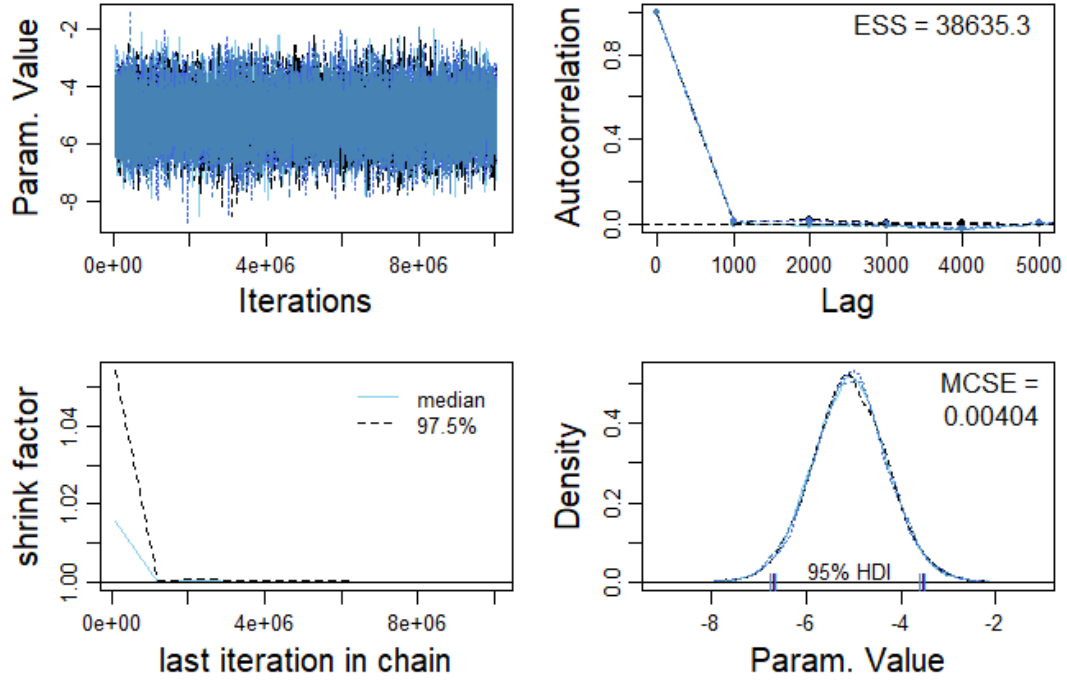


Fig. A14. Posterior of m_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

S.C

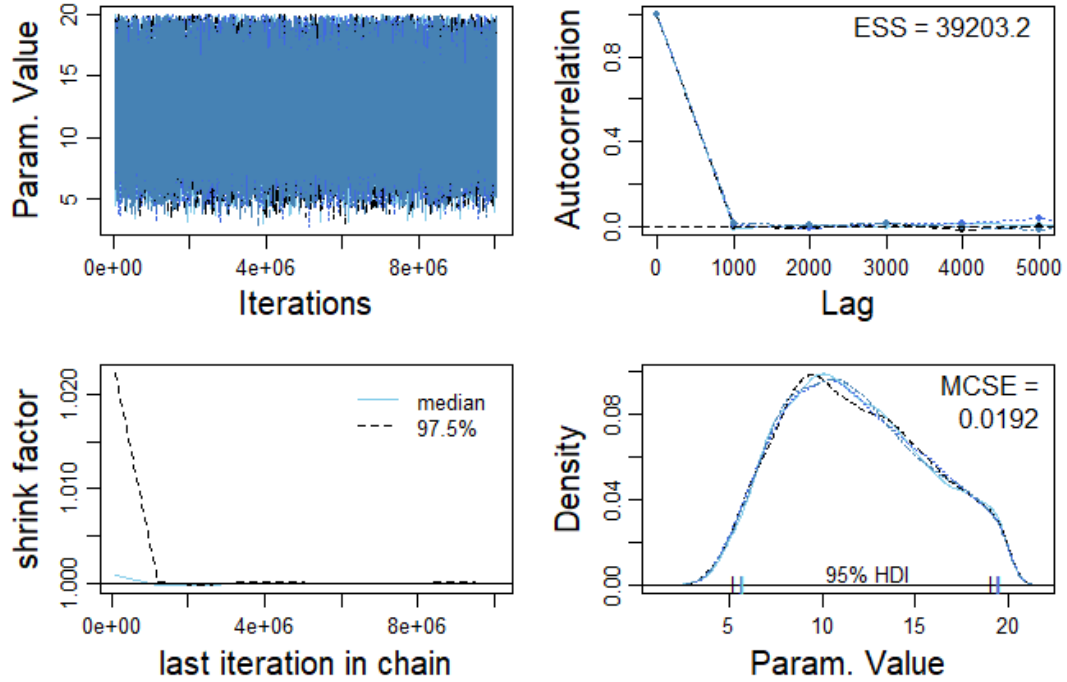


Fig. A15. Posterior of s_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

d.c

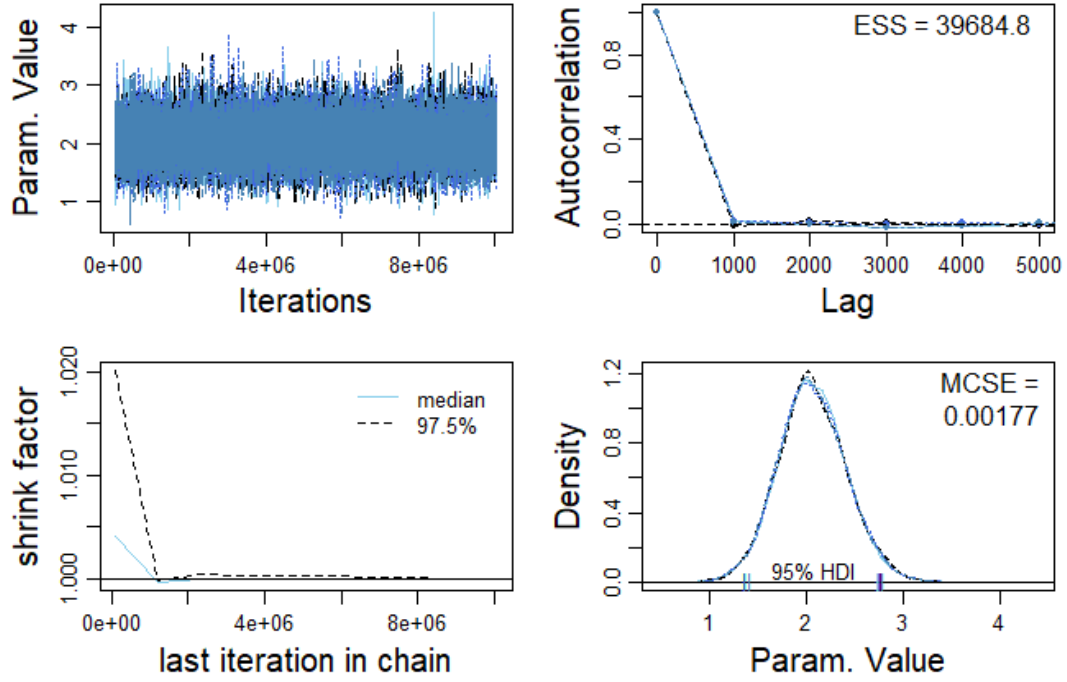


Fig. A16. Posterior of d_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.hc

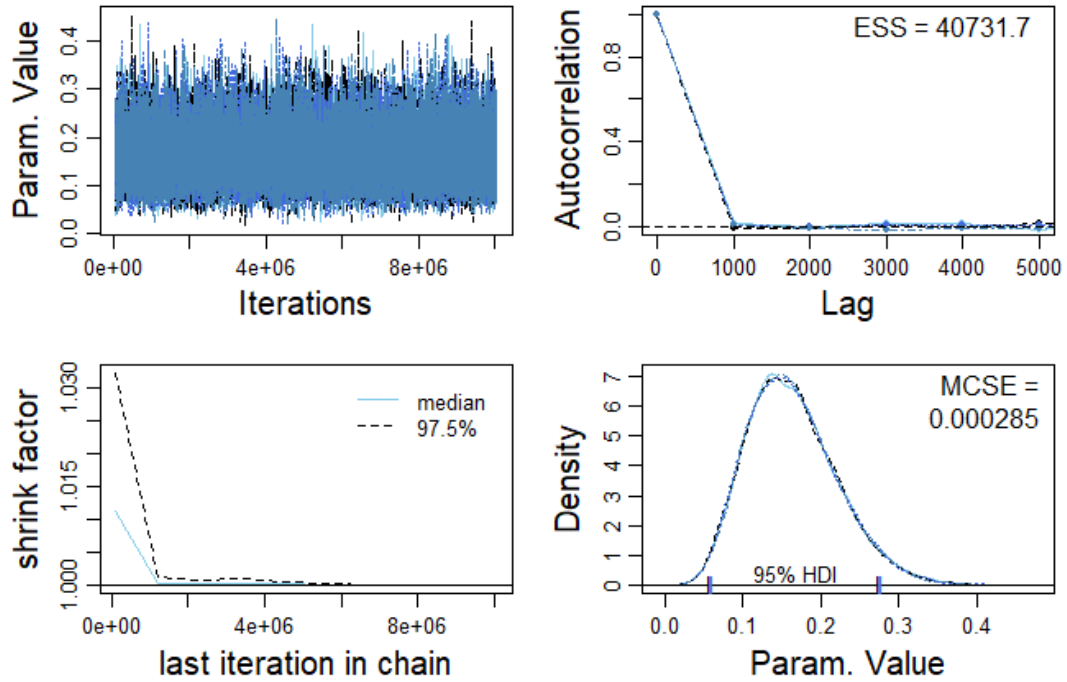


Fig. A17. Posterior of p_{hc} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.hh

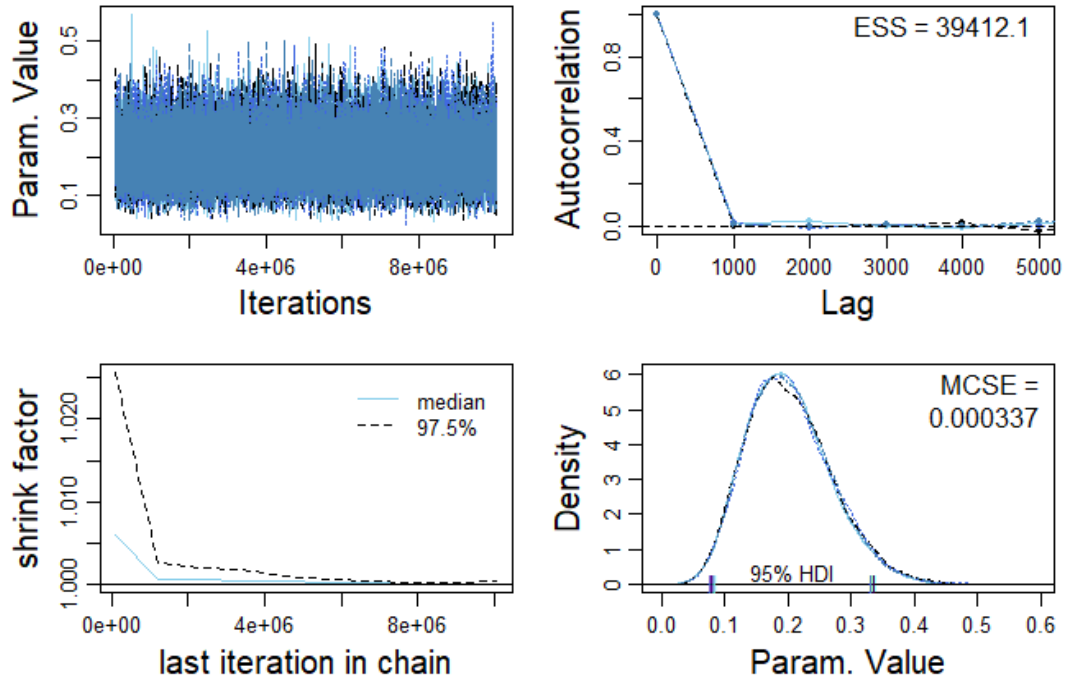


Fig. A18. Posterior of p_{hh} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.h

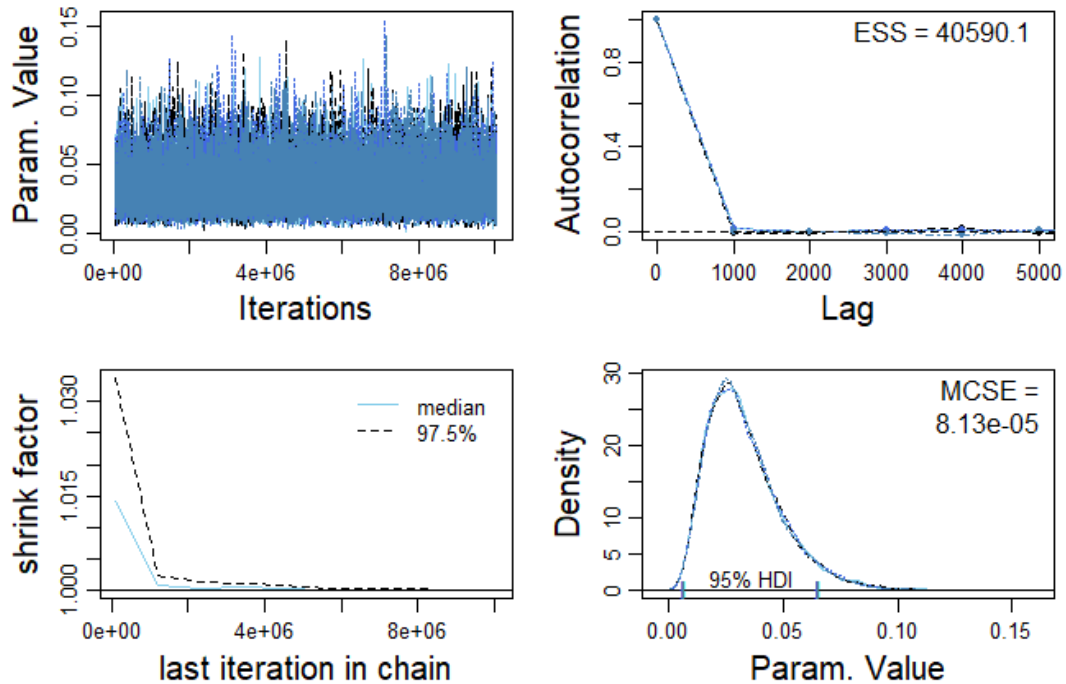


Fig. A19. Posterior of p_h . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

lambda.c

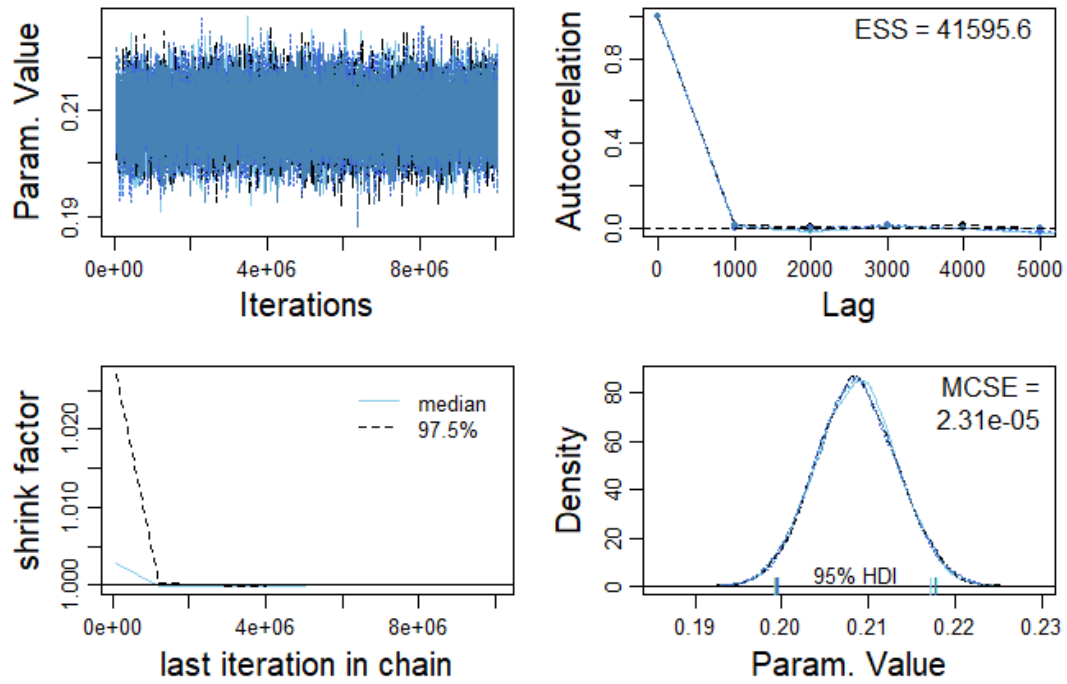


Fig. A20. Posterior of λ_c . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

lambda.e

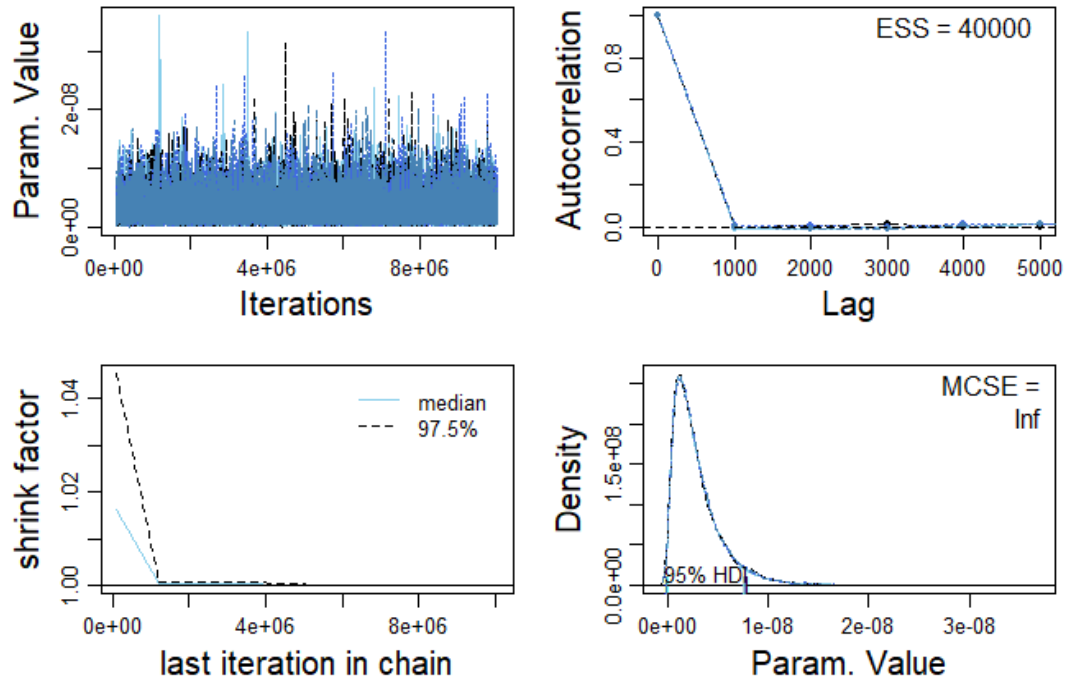


Fig. A21. Posterior of λ_e . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run. The MCSE of 'Inf' indicates a value below machine precision.

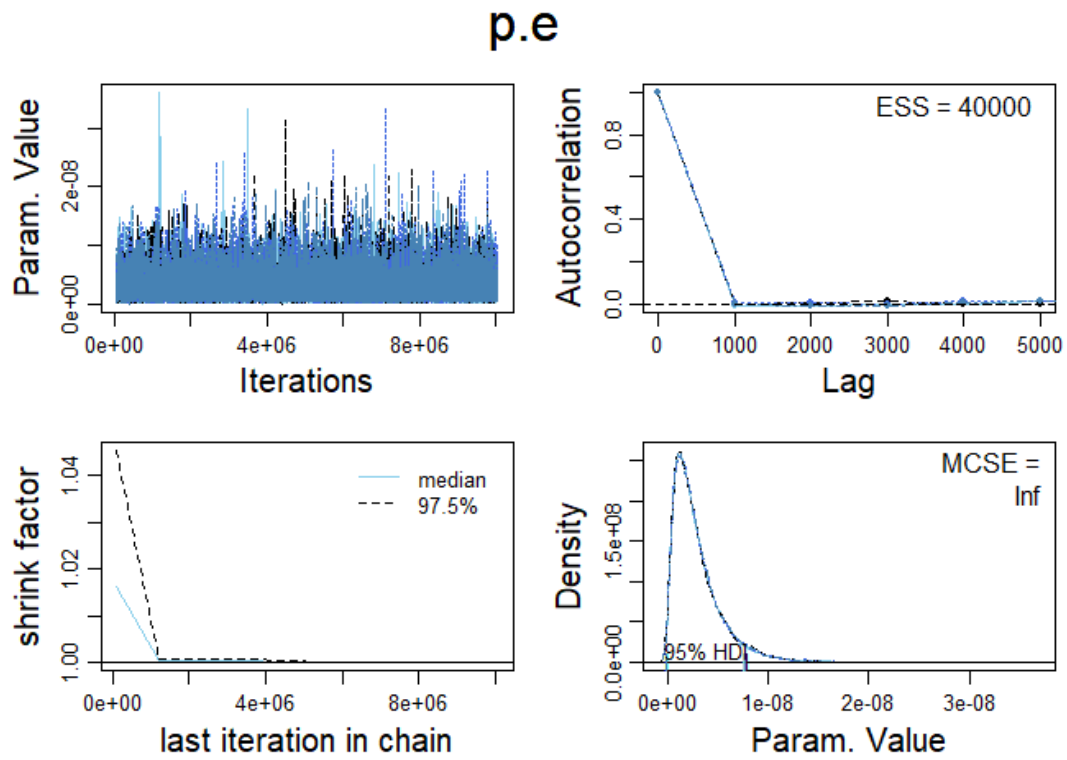


Fig. A22. Posterior of p_e . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run. The MCSE of ‘Inf’ indicates a value below machine precision.

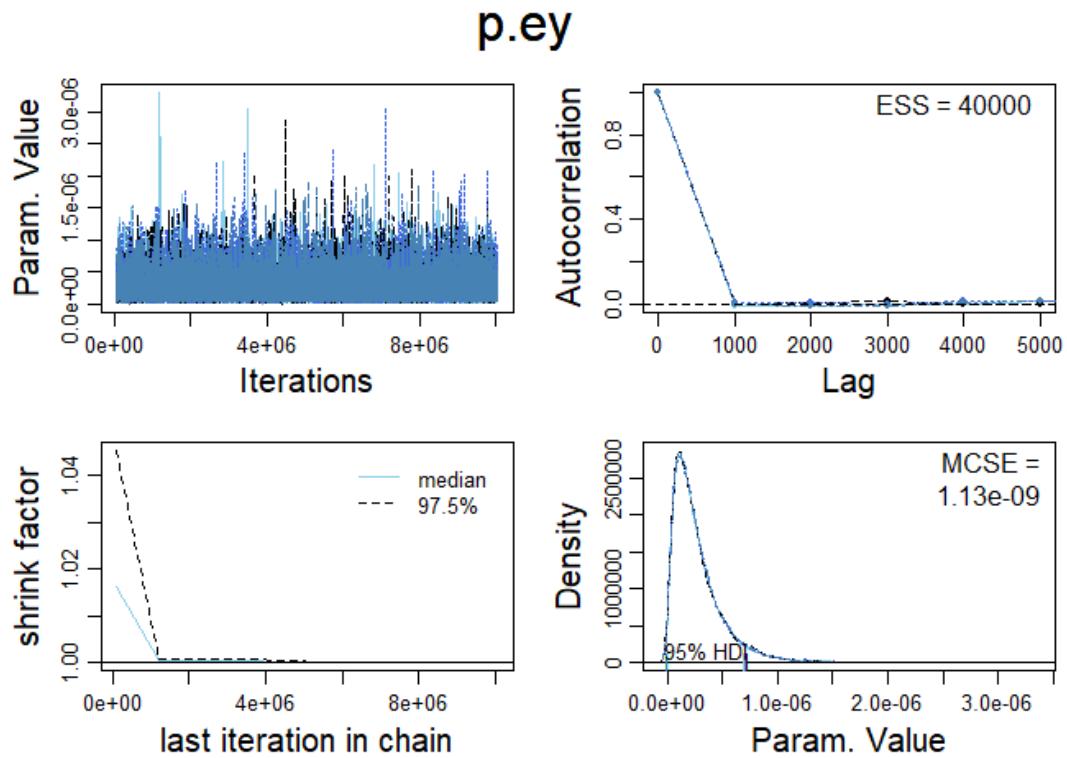


Fig. A23. Posterior of p_{ey} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.is

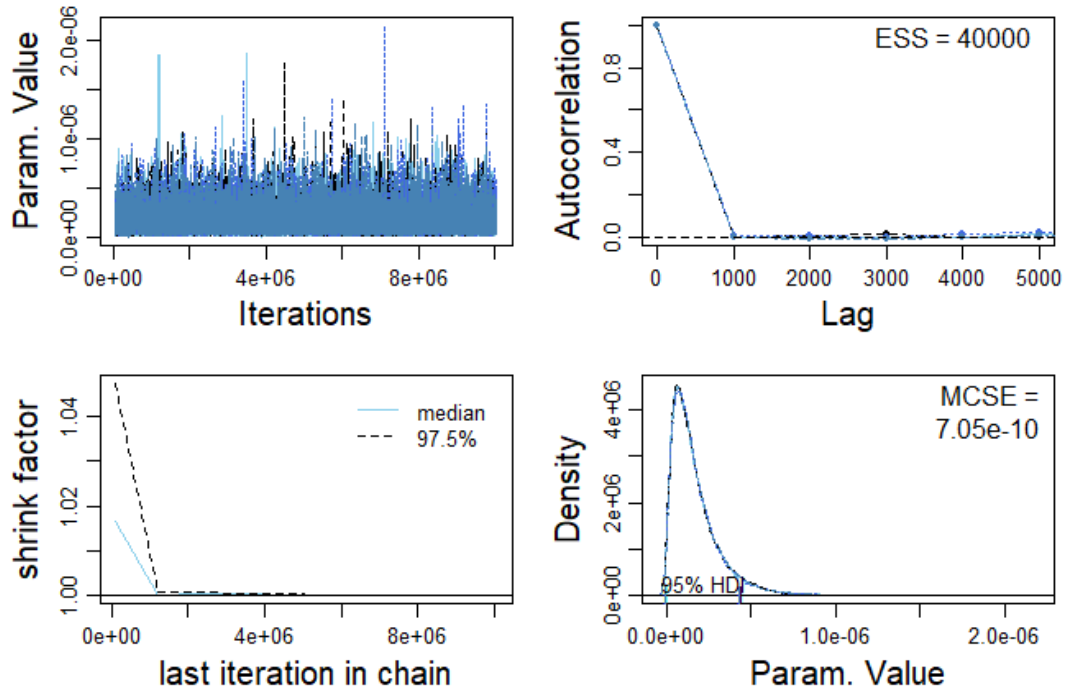


Fig. A24. Posterior of p_{is} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p.ie

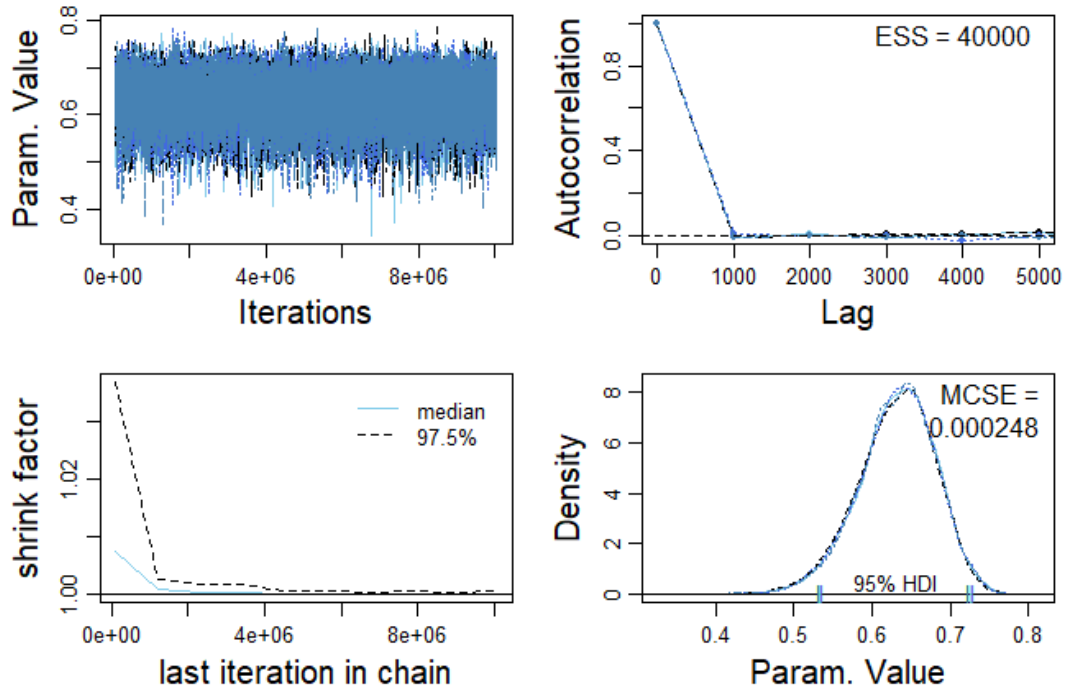


Fig. A25. Posterior of p_{ie} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

D

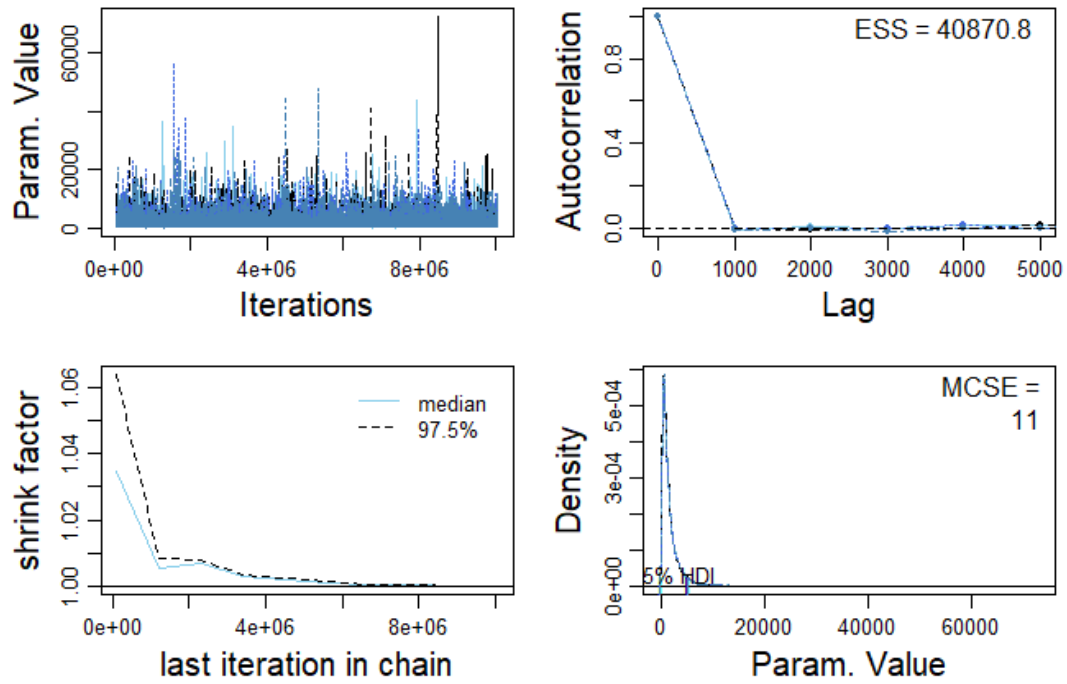


Fig. A28. Posterior of d . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

p_{it}

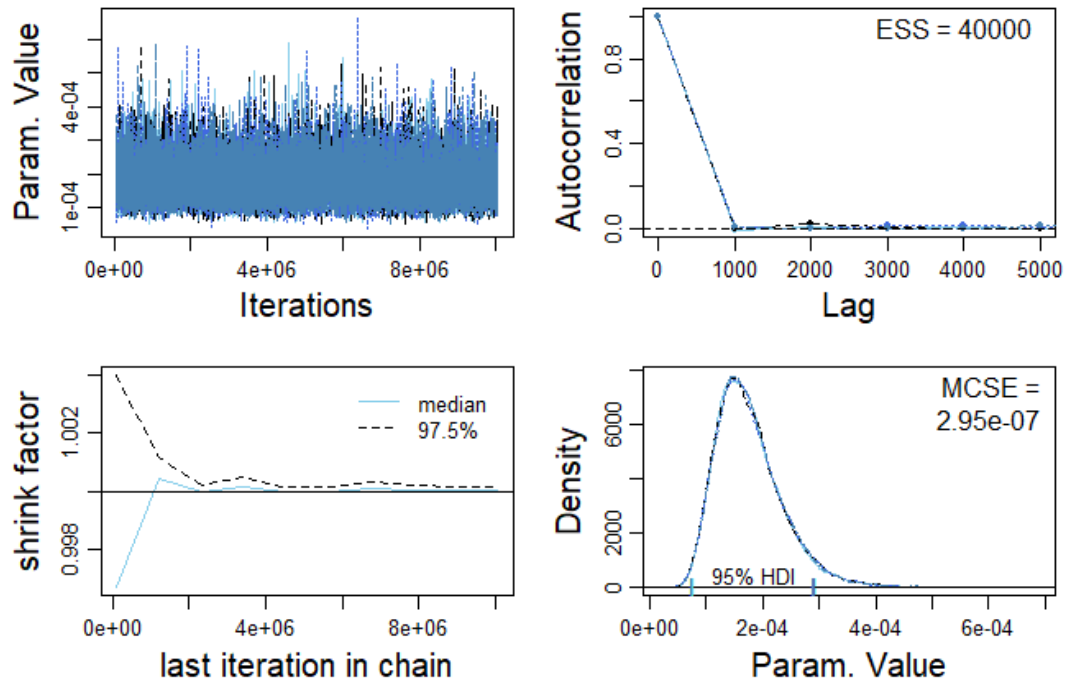


Fig. A29. Posterior of p_{it} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.

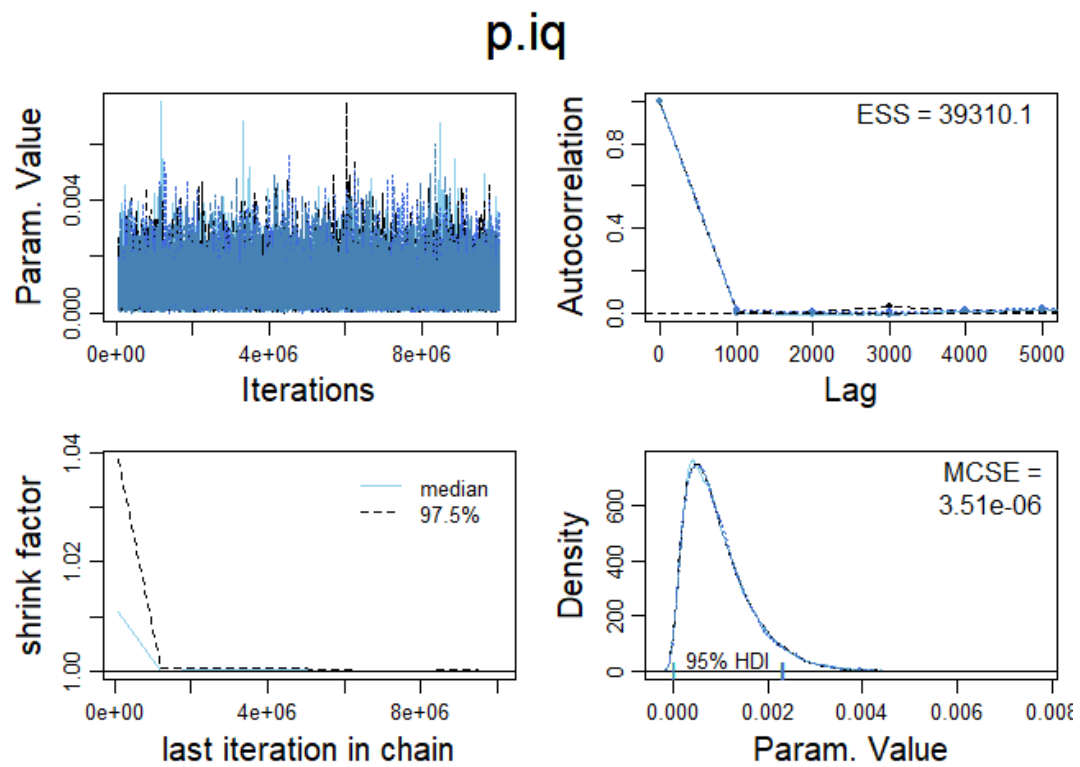


Fig. A30. Posterior of p_{iq} . Clockwise from top-left: A trace plot of the chain trajectories. A measure of the chain trajectory autocorrelation. A density plot of the posterior distribution for both chains, with 95% HDI intervals marked and MCSE provided. A plot of chain shrink factor across the model run.