COMPUTO
ISSN 2824-7795

A hierarchical model to evaluate pest treatments from prevalence and intensity data

Armand Favrot[1*] **David Makowski**[1†]

[1]MIA Paris-Saclay, INRAE AgroParisTech Université Paris-Saclay, France

Date published: 2024-01-09 Last modified: 2026-10-18

DOI: 10.57750/6cgk-g727

Abstract

In plant epidemiology, pest abundance is measured in field trials using metrics assessing either pest prevalence (fraction of the plant population infected) or pest intensity (average number of pest individuals present in infected plants). Some of these trials rely on prevalence, while others rely on intensity, depending on the protocols. In this paper, we present a hierarchical Bayesian model able to handle both types of data. In this model, the intensity and prevalence variables are derived from a latent variable representing the number of pest individuals on each host individual, assumed to follow a Poisson distribution. Effects of pest treaments, time trend, and between-trial variability are described using fixed and random effects. We apply the model to a real data set in the context of aphid control in sugar beet fields. In this data set, prevalence and intensity were derived from aphid counts observed on either factorial trials testing different types of pesticides treatments or field surveys monitoring aphid abundance. Next, we perform simulations to assess the impacts of using either prevalence or intensity data, or both types of data simultaneously, on the accuracy of the model parameter estimates and on the ranking of pesticide treatment efficacy. Our results show that, when pest prevalence and pest intensity data are collected separately in different trials, the model parameters are more accurately estimated using both types of trials than using one type of trials only. When prevalence data are collected in all trials and intensity data are collected in a subset of trials, estimations and pest treatment ranking are more accurate using both types of data than using prevalence data only. When only one type of observation can be collected in a pest survey or in an experimental trial, our analysis indicates that it is better to collect intensity data than prevalence data when all or most of the plants are expected to be infested, but that both types of data lead to similar results when the level of infestation is low to moderate. Finally, our simulations show that it is unlikely to obtain accurate results with fewer than 40 trials when assessing the efficacy of pest control treatments based on prevalence and intensity data. Because of its flexibility, our model can be used to evaluate and rank the efficacy of pest treatments using either prevalence or intensity data, or both types of data simultaneously. As it can be easily implemented using standard Bayesian packages, we hope that it will be useful to agronomists, plant pathologists, and applied statisticians to analyze pest surveys and field experiments conducted to assess the efficacy of pest treatments.

Keywords: bayesian model, epidemiology, hierarchical model, pest control, trial, survey

*Corresponding author: armand.favrot@inrae.fr
†Corresponding author: david.makowski@inrae.fr

Contents

1 Introduction **2**

2 Material and Method **4**
2.1 Description of the data 4
2.2 Model 4
2.2.1 Specification 4
2.2.2 Inference on real data 7
2.2.3 Simulations 10

3 Results **16**
3.1 Results obtained with real data 16
3.1.1 Model convergence and posterior distributions 16
3.1.2 Estimated values of pesticide treatment efficacies 19
3.2 Results obtained by simulation 20
3.2.1 Interest of combining trials with prevalence and trials with intensity 20
3.2.2 Interest of adding intensity when prevalence is measured in all trials 22
3.2.3 Is it better to measure intensity or prevalence in new pest surveys? 23

4 Conclusion **24**

5 Author contributions **24**

6 Funding **24**

7 Data availability **26**

8 Acknowledgements **26**

Supplementary material **26**
Model adjustment and comparison to the negative binomial model 26

References **26**

1 Introduction

In plant epidemiology, pest and disease presence can be measured in a host population using different metrics. A first metric measures the presence/absence of the pest in the individuals (plants) of the host population. This metric is often called prevalence or incidence (Madden and Hughes (1999), Shaw et al. (2018)). The prevalence describes the proportion of the host population in which the pest is present. This metric is relevant and widely used, but it does not account for the number of pest individuals per host individual. With the prevalence, a plant infected by one single pest individual (e.g., an insect) and a plant infected by many pest individuals both represents one infected plant. For this reason, pest abundance is sometimes assessed using another metric representing the average number of pest individuals per host individual. This metric is called intensity or severity (Madden and Hughes (1999), Shaw et al. (2018)), and describes the intensity of the disease in the target population. These two metrics do not generally have the same requirement in terms of working time; measuring intensity takes indeed much more time than measuring prevalence because it is very tedious to count all pest individuals, especially when individuals are small, numerous, and/or difficult to detect.

Pest prevalence and intensity are commonly measured in factorial field trials to test the efficacy

of different treatments. In this paper, we place ourselves in an important application framework which is the evaluation of alternative pesticide treatments to neonicotinoids against aphids in sugar beet. Indeed, neonicotinoids had been a popular chemical treatment to control aphids for many years, especially in sugar beets, a major crop in Europe. Recently, neonicotinoids were recognized as presenting high risks for the environment with a negative impact on a wide range of non-target organisms, including bees (Wood and Goulson (2017), Pisa et al. (2015)), and this family of pesticides has been banned in several European countries. In order to find a substitute to neonicotinoids, a number of factorial field trials were conducted to compare the efficacy of different alternative treatments during several years in different countries. Each trial consists of a set of plots divided into several blocks, themselves divided into several strips on which different pesticide treatments are randomly allocated. One strip always remains untreated to serve as a control. In each strip, aphid prevalence, aphid intensity or both metrics are measured in a sample of plants (usually, 10-20 plants per strip). Depending on the protocol and on the working time constraint, either one type or both types of metrics are measured. Consequently, for a given pest treatment, some trials may report prevalence data while others report intensity data or both types of data. This heterogeneity raises several issues. A first issue concerns the statistical analysis of the trials reporting prevalence and intensity. Although it is easy to fit a generalized linear model to each type of data separately, it is less straightforward to fit a single model to the whole set of trials in order to obtain a single ranking of the pest treatments taking into account both types of data at the same time. Generally, factorial trials assessing treatment efficacy are analyzed with statistical models that take into account one of the two metrics but not both. Prevalence data are thus commonly analyzed using binomial generalized linear models and intensity data are frequently analyzed with Poisson generalized linear models (Michel et al. (2017), Laurent et al. (2023), Agresti (2015)). As far as we know, no statistical model has been proposed to assess treatment efficacy based on the simultaneous analysis of prevalence and intensity data. In Osiewalski and Marzec (2019), the authors introduced a switching model designed to handle two count variables, one of which may be degenerate. This model was employed to characterize the counts of cash payments and bank card payments in Poland, utilizing data from both cardholders and non-cardholders. A generalized form of the bivariate negative binomial regression model was developed in Gurmu and Elder (2000), allowing for a more flexible representation of the correlation between the dependent variables. This model was applied to describe the number of visits to a doctor and the number of visits to non-doctor health professionals. It outperformed existing bivariate models across various model comparison criteria. In order to analyze data related to crash counts categorized by severity, Park and Lord (2007) employed a multivariate Poisson log-normal model, effectively addressing both over-dispersion and a fully generalized correlation structure within the data set. However, it should be noted that these models did not include any binomial distribution and thus could not be used to deal with proportion data, such as pest prevalence. Another issue concerns the practical value of combining both prevalence and intensity data. It is unclear whether the simultaneous analysis of prevalence and intensity data may increase the accuracy of the estimated treatment efficacy compared to the use of a single type of data, and whether this may increase the probability of identifying the most efficient treatments. Finally, it is also unclear how future trials should be designed, in particular how many trials are required to obtain accurate estimations, and whether intensity data should be preferred to prevalence data.

In this paper, we propose a new flexible statistical model that can be used to rank pest treatments from trials including prevalence data, intensity data, or both. We apply it to a real data set including trials testing the efficacy of pesticides against aphids infesting sugar beets, considering contrasted scenarios of data availability, and we show how the proposed model can be used to evaluate the efficacy of different treatments. Based on simulations, we then quantify the reduction of mean absolute errors in the estimated treatment efficacy resulting from the use of both prevalence and intensity data during the statistical inference, compared to the use of either prevalence or intensity data. The rest

of the paper is organized as follows. First, we present the structure of the data set including real prevalence and intensity data. Next, we describe in detail the proposed model, the inference method, and the simulation strategy. After checking the convergence of the fitting algorithm, we show how the model can be used to assess treatment efficacy. We finally present the results based on simulated data and we make several recommendations.

2 Material and Method

2.1 Description of the data

Data are collected in 32 field trials conducted in France, Belgium and the Netherlands to compare several treatments against aphids in sugar beets. Each trial consists in a plot located in a given site at a given year (site-year) divided into one to four blocks. Each of these blocks is itself divided into strips where different treatments are tested, one of these treatments being an untreated control and the others corresponding to different types of insecticide. In each strip of each block, the number of aphids is counted on a sample of 10 beet plants (intensity). The number of infested plants (prevalence) is measured as well, but only in 15 trials out of 32. The total numbers of intensity and prevalence data are equal to 1128 and 561, respectively. Note that the number of aphids is not counted on each beet plant but in the whole plant sample. Intensity and prevalence are monitored at different times after treatments. As shown in Figure 1 A, the data set is unbalanced as less data are available for the treatment Mavrik-jet compared to the others. Figure 1 B shows that the intensity and prevalence tend to increase with time.

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```

```
Warning in get_plot_component(plot, "guide-box"): Multiple components found;
returning the first one. To return all, use `return_all = TRUE`.
```

2.2 Model

2.2.1 Specification

We introduce an unobserved variable representing the number of pest individuals (here, aphids) on each plant in a sample of N plants (here, sugar beets). This variable is noted W and is assumed to follow a Poisson probability distribution whose mean value is a function of time.

We use the following indexes: i for the trial, j for the treatment, k for the block, t for the time and s for the plant number. The distribution of W_{ijkts} is defined as:

$$W_{ijkts} \sim \mathscr{P}(\lambda_{ijkt}) \tag{1}$$

$$\log\ \lambda_{ijkt} = \alpha_0 + \beta_{0i} + \gamma_{0j} + (\alpha_1 + \gamma_{1j})\, X_t + u_{ij} + \epsilon_{ijkt} \tag{2}$$

with

- $\beta_{0i} \sim \mathcal{N}(0, \sigma_0^2)$
- $u_{ij} \sim \mathcal{N}(0, \chi^2)$
- $\epsilon_{ijkt} \sim \mathcal{N}(0, \eta^2)$

The random variables are all assumed independent. The parameters α_0, α_1, γ_{0j}, and γ_{1j} are considered as fixed. This model serves as a tool for conducting inference on a population of trials, from which the

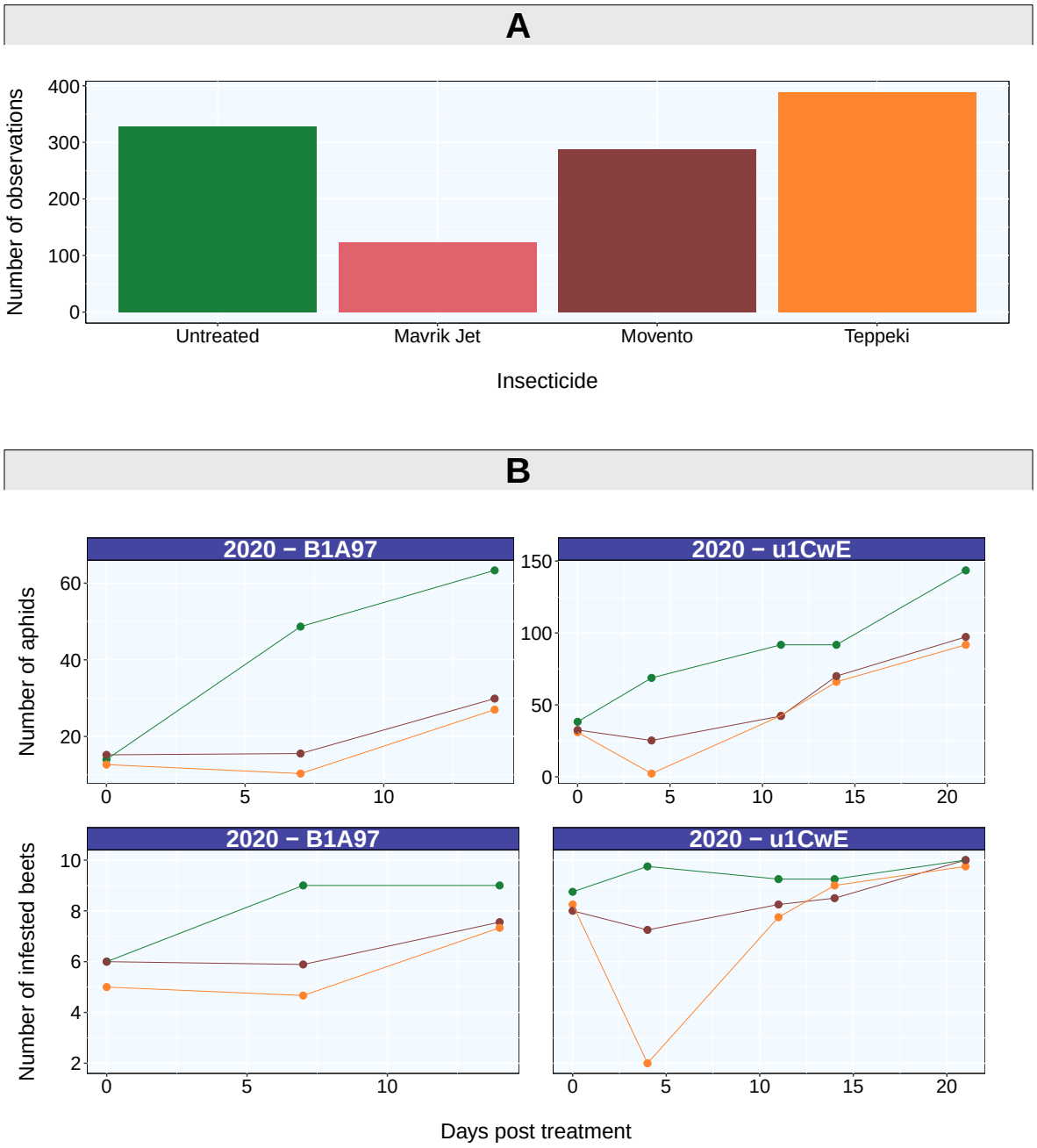

Figure 1: Description of the data set. **A** Number of observations according to the type of insecticide. **B** Examples of observed number of aphids averaged over the blocks (intensity) and number of infested beets out of ten (prevalence) averaged over the blocks, at different dates for two trials.

subset of trials comprising our data set is assumed to constitute a random sample. In essence, the trials contained within our database are leveraged to estimate the parameter values that characterize a target population, where the tested pest control treatments will be actually implemented. Consequently, all parameters contingent on individual trials have been defined as random effects.

The observed variables (intensity and prevalence) can be expressed as a function of W. We note:

- Y_{ijkt} the number of pest individuals (aphids) in the sample of N_i plants collected in trial i, treatment j, block k, at time t
- Z_{ijkt} the number of infested plants among the N_i plants collected in trial i, treatment j, block k, at time t

Then, assuming the Ws are independent, we have:

$$Y_{ijkt} = \sum_{s=1}^{N_i} W_{ijkts} \qquad Y_{ijkt} \sim \mathscr{P}(N_i \lambda_{ijkt}) \tag{3}$$

$$Z_{ijkt} = \sum_{s=1}^{N_i} 1_{W_{ijkts}>0} \qquad Z_{ijkt} \sim \mathscr{B}(N_i,\ \pi_{ijkt}) \tag{4}$$

where $\pi_{ijkt} = 1 - \exp(-\lambda_{ijkt})$. The different quantities used in the model are defined in Table 1.

Table 1: Description of the indices, inputs and parameters used in the model

i	trial index
j	treatment index
k	block index
t	time index
s	plant index
N_i	sample size (number of plants)
λ_{ijkt}	mean number of pest individuals (aphids) on one plant
π_{ijkt}	probability for a plant to be infested
α_0	mean number of pest individuals (aphids) in the untreated group
β_{0i}	trial effect
γ_{0j}	effect of treatment j at time 0 (time of treatment)
α_1	growth parameter of the number of pest individuals for the untreated group
γ_{1j}	effect of the treatment on the time effect (interaction between treatment j and time)
X_t	number of days post treatment
u_{ij}	random interaction between trial and treatment
ϵ_{ijkt}	residuals

From this model we define the efficacy of the jth treatment at time t (t days after pesticide application) by the quantity (Laurent et al. (2023)):

$$\mathrm{Ef}_{jt} = \Big(1 - \exp\big(\gamma_{0j} + \gamma_{1j} \times X_t\big)\Big) \times 100 \tag{5}$$

The quantity Ef_{jt} corresponds to the expected percentage reduction of pest individuals (aphid numbers) for the j-th treatment compared to the untreated group, over trials and blocks.

Our Poisson log linear model includes an additive random dispersion term associated to each individual observation (ϵ_{ijkt} in Equation 2). This is a standard and well-recognized approach to deal with over-dispersion (Harrison (2014)). In order to check the model assumptions, we perform a posterior predictive check of our model to check that the data were compatible with the model assumptions (Supplementary material). Posterior predictive check is frequently used to look for systematic discrepancies between real and simulated data (Gelman et al. (1995)). To do so, we compute the probability of exceeding each individual data with the fitted model (Equation 2). The computed probabilities are all falling in the range 0.22-0.93 (except for the observations equal to 0, for which the probability of being greater was equal to 1), and are thus not extreme. This result indicates that the model specified is not incompatible with the observed data and that the over-dispersion is correctly taken into account. In addition, we fit another model including a negative binomial distribution instead of a Poisson distribution. The results are almost identical between the two models (Supplementary material).

2.2.2 Inference on real data

The model parameters are estimated using Bayesian inference with a Markov chain Monte-Carlo method. We perform the inference using R, with the package rjags (Plummer (2022)). For each of the data sets listed in Table 2, we fit the model (Equation 2 - Equation 4) with the following weakly informative priors: $\mathcal{N}(0, 10^3)$ for the parameters α_0, γ_0, α_1, γ_1 and $\mathcal{U}([0, 10])$ for the parameters σ_0, χ, η. We use two Markov chains with 2×10^5 iterations (after an adaptation phase of 10^5 iterations), and we center the time variable t to facilitate convergence.

The convergence of the MCMC algorithm is checked by inspecting the mixing of the two Markov chains and monitoring the Gelman-Rubin diagnosis statistics (Gelman and Rubin (1992)). We then compute the posterior mean of the pesticide treatment efficacy (defined by Equation 5) as well as its 95% credibility interval. The code used to fit the model is provided below.

> **Note**
>
> The following code presents the inference on an extract of the real data set, which includes both trials of figure 1B (2020 - B1A97 ; 2020 - u1CwE). It is a demo for the "50% Y - 50% Z" scenario and we set here the number of adaptation and iteration to 2000 in order to reduce computation time.

```
# Jags code for the model  ##################################################
modelstringYZ = "
    model {

        # Likelihood ########################################################
        for (i in 1:Q){
            Y[i] ~ dpois(N[i] * lb[i])
            Z[i] ~ dbinom(pi[i], N[i])

            log(lb[i]) = beta0[ID[i]] + gamma0[INSEC[i]] + (alpha1 +
                         gamma1[INSEC[i]]) * TIME[i] + u[ST[i]] + epsi[i]

            pi[i] = 1 - exp(- lb[i])
            epsi[i] ~ dnorm(0, pi_eps)
        }
```

```
        for (j in 1:K){
            beta0[j] ~ dnorm(alpha0, tau0)
        }

        for (c in 1:M){
            u[c] ~ dnorm(0, invchi)
        }

        gamma0[1] = 0
        gamma1[1] = 0

        # Priors ##########################################################
        for (s in 2:L){
            gamma0[s] ~ dnorm(0, 0.001)
            gamma1[s] ~ dnorm(0, 0.001)
        }

        alpha0 ~ dnorm(0, 0.001)
        alpha1 ~ dnorm(0, 0.001)
        sigma0 ~ dunif(0, 10)
        chi ~ dunif(0, 10)
        eta ~ dunif(0, 10)

        # Derived Quantities ##############################################
        tau0 = pow(sigma0, -2)
        invchi = pow(chi, -2)
        pi_eps = pow(eta, -2)

        for (h in 2:L){
            for(t in 1 : T){
                Eff[h, t]  = (1 - exp(gamma0[h] + gamma1[h] *
                              TIME_unique[t])) * 100
            }
        }
    }
"

writeLines(modelstringYZ, con  =  "jags_models/modelYZ.txt")
##########################################################################


# Inference example on the extract of the real dataset #####################
real_data_extract = readRDS(file = "data/real_data_extract.rds")

data = real_data_extract %>%
  mutate(tscaled = scale(DPT), st = paste(ID, Insecticide))

# Building scenarios -------------------------------------------------------
scenarioY = data %>% mutate(Z = NA)
scenarioYhalfZhalf = data %>% mutate(Y = ifelse(ID == "2020 - B1A97", Y, NA),
```

```
                                    Z = ifelse(ID == "2020 - B1A97", NA, Z))
scenarioYhalf = data %>% filter(ID == "2020 - B1A97") %>% mutate(Z = NA)
scenarioZhalf = data %>% filter(ID == "2020 - u1CwE") %>% mutate(Y = NA)
# ---------------------------------------------------------------------------

data = scenarioYhalfZhalf

Y = data$Y;   Q = length(Y);   N = data$N;   Z = data$Z

ID = as.numeric(as.factor(as.character(data$ID)));
INSEC = as.numeric(as.factor(as.character(data$Insecticide)));
TIME = as.numeric(data$tscaled);
ST = as.numeric(as.factor(as.character(data$st)));

K = length(unique(ID));   L = length(unique(INSEC));
M = length(unique(ST));

df_TIME = suppressMessages(data %>%
                             group_by(DPT, tscaled) %>%
                             summarise(n = n()) %>% as.data.frame)

TIME_unique = approx(df_TIME$DPT, df_TIME$tscaled, xout = c(6, 12))$y;
T = length(unique(TIME_unique))

data_jags = list(
  "Y" = Y, "Z" = Z, "Q" =  Q, "ID" = ID, "INSEC" = INSEC,
  "TIME" = TIME, "ST" = ST, "K" = K, "L" = L, "M" = M,
  "N" = N, "T" = T, "TIME_unique" = TIME_unique
)

nadapt = 2000;   niter = 2000

model <- jags.model("jags_models/modelYZ.txt", data = data_jags,
                    n.chains = 2, n.adapt = nadapt)

samples <- coda.samples(model,
                        variable.names = c("gamma0", "gamma1", "Eff"),
                        n.iter = niter, thin = 10)
###########################################################################
```

In practice, it is common that only Y or Z data are available in some of the trials. In this case, the resulting data set includes observations of Y in some trials and observations of Z in others. The data set may even include one type of observations only, either Y or Z, in all trials. Here, we define four scenarios with contrasted levels of Y and Z availability in order to evaluate the consequences of using different types of data sets. We consider four data subsets defined from the real data set including trials with observations of Y, with observations of Z, or with both types of observation in different proportions (Table 2). The data subset "100% Y - 0% Z" includes Y data collected in the 32 trials. The data subset "50% Y - 0% Z" includes Y data collected in the 17 trials for which no Z observation is available. The data subset "0% Y - 50% Z" includes the Z data collected in the 15 trials

for which Z observations are available. The data subset "50% Y - 50% Z" includes Y data collected in 17 trials and Z data collected in the other 15 trials. The latter data subset does not include any trial reporting both Y and Z data. Throughout our analysis, missing data are assumed to be missing at random.

The hierarchical model defined above is fitted to each data set in turn. Each fitted model is then used to compute the posterior mean and 95% credibility interval of Ef_{jt} for each treatment at t=6 and 12 days after pesticide application.

Table 2: Four data subsets defined from the original data set (real data).

Type of data set	Description
100% Y - 0% Z	Y observations available in the 32 trials and no Z
50% Y - 0% Z	Y observations available in 17 trials and no Z
0% Y - 50% Z	Z observations available in 15 trials and no Y
50% Y - 50% Z	Y observations available in 17 trials and Z observations available in the other 15 trials

2.2.3 Simulations

Simulations are carried out to further investigate the impact of the type and amount of data available on the accuracy of the parameters and the ability of the model to identify the most and least effective treatments.

We define three numbers of trials, equal to 20, 40 and 80, successively. For data simulations, the model parameters are set equal to those estimated from the real data set "100% Y - 0% Z" defined in Table 2 (posterior means). For each number of trials, we generate virtual data from the model (Equation 1 - Equation 4) and estimate the model parameters, according to the following procedure:

- Draw values of β_{0i}, u_{ij} and ϵ_{ijkt} in their distributions for each trial, 3 treatments (+ the untreated control), 3 dates (t=0, 6, 12), and 4 blocks
- Calculate λ_{ijkt} from Equation 2,
- Draw values of W_{ijkts} in its Poisson distribution for 10 plants (s=1, ..., 10)
- Calculate Y_{ijkt} and Z_{ijkt} from the Ws for each trial, treatment, date, block.
- Generate the eight data subsets corresponding to the scenarios defined in Table 3 (including all values of W, the generated values of Y only, the generated values of Z only, both Y and Z values but not W, the values of Y in 50% of the trials, the values of Z in 50% of the trials, the values of Y in 50% of the trials and the values of Z in the other 50%),
- Fit the model (Equation 1 - Equation 4) to each of the data subsets according to the procedure described above based on MCMC.

At the end of this procedure, we get eight sets of estimated parameters, corresponding to the eight scenarios defined in Table 3. This procedure is repeated 1000 times with each time a different seed between 0 and 999. However, the computations performed with jags failed for 26 replicates and thus 974 replicates were available for the analysis.

For each number of trials and each scenario defined in Table 3, the accuracy of the estimated parameters γ (from which depend the treatment efficacies) is evaluated by computing an absolute error, averaged over the three treatments (j=1 corresponding to the control) as:

$$E_\gamma = \frac{1}{2 \times 3} \sum_{j=2}^{4} \left(\frac{|\gamma_{0j} - \hat{\gamma}_{0j}|}{|\gamma_{0j}|} + \frac{|\gamma_{1j} - \hat{\gamma}_{1j}|}{|\gamma_{1j}|} \right) \tag{6}$$

where the true parameter values are set equal to the posterior means computed with the real data set, and the parameter estimates ($\hat{\gamma}_{0j}$ and $\hat{\gamma}_{1j}$) are the posterior means computed for the j-th treatment. For each trial number and each scenario defined in Table 3, the 974 values of E_γ obtained for the 974 generated data subsets are then averaged. The average values obtained for the eight scenarios are finally compared to determine the type of data leading to the most accurate estimated parameter values.

In addition, we compare the eight scenarios according to another criterion measuring the difference between the estimated efficacy values and the true efficacy value (averaged over the three pesticide treatments considered), as follows:

$$E_{\mathrm{Ef}_t} = \frac{1}{3}\sum_{j=2}^{4}\frac{|\mathrm{Ef}_{jt} - \hat{\mathrm{Ef}}_{jt}|}{|\mathrm{Ef}_{jt}|} \tag{7}$$

where the true treatment efficacy is defined by Equation 5 (setting the parameters γs to the posterior means obtained with the real data set) and the estimated efficacy ($\hat{\mathrm{Ef}}_{jt}$) is the posterior mean computed with the simulated data set. The 974 values of E_{Eft} obtained from the 974 simulated data sets are then averaged for each trial number and each scenario. Finally, we evaluate the proportions of cases where the true best treatment (i.e., the treatment with the highest efficacy) is correctly identified.

In order to determine whether the difference of performance obtained with the types of data Y and Z depends on the pest abundance, we perform an additional series of simulations with three values for the model parameter α_0, equal to -1, 1 and 2, successively. These three values of α_0 define three contrasted levels of pest abundance (the higher α_0, the higher the abundance). We used the procedure outlined above considering only two scenarios of data availability, namely "100% Y - 0% Z" and "0% Y - 100% Z". This procedure is implemented with each value of α_0 in turn. The results are used to compare the model performances using either Y or Z for parameter estimation, depending of the pest abundance specified by α_0.

Table 3: Eight scenarios compared using simulated data.

Type of data set	Description
100% W	W observations available in all the trials
100% Y - 0% Z	Y observations available in all the trials
0% Y - 100% Z	Z observations available in all the trials
100% Y - 100% Z	Y and Z observations available in all the trials
50% Y - 0% Z	Y observations available in half of the trials
0% Y - 50% Z	Z observations available in half of the trials
50% Y - 50% Z	Y observations available in half of the trials and Z observations available in the other half of the trials
50% Y - 100% Z	Y observations available in half of the trials and Z observations available in all the trials

These simulations required a significant amount of computation time and were conducted on the INRAE MIGALE server. For a single seed and 20 trials, the computations took 80 minutes on a computer with an Intel Core i7 processor running at 1.90 GHz and 32 GB of RAM. The code used to perform them is presented below.

```
simu_data <- function(seed, I){
```

```
set.seed(seed)

ID = c(1 : I);

data = expand_grid(ID = c(1 : I),   Block = c(1 : K),   Band = c(1 : J),
                   Beet = c(1 : N),   DPT = seq(0, 12, length.out = T))

data = data %>%
 mutate(Insecticide = ifelse(DPT <= 0, "T 1", paste("T", Band)),
        tscaled = scale(data$DPT)) %>%
 mutate(st = paste(ID, Insecticide),
        sbit = paste(ID, Block, Insecticide, DPT))

Insecticide = data$Insecticide %>% unique %>% sort
names(gamma0) = Insecticide
names(gamma1) = Insecticide

beta0 = rnorm(I, sd = sig0);
u = rnorm(I * J, sd = chi);
epsi = rnorm(I * J * K * (T - 1) + I * K, sd = eta);

names(beta0) = data$ID %>% unique;
names(u) = data$st %>% unique
names(epsi) = data$sbit %>% unique

data = data %>% mutate(alpha0 = alpha0,
                       alpha1 = alpha1,
                       N = N,
                       beta0 = recode(ID, !!!beta0),
                       gamma0 = recode(Insecticide, !!!gamma0),
                       gamma1 = recode(Insecticide, !!!gamma1),
                       u = recode(st, !!!u),
                       epsi = recode(sbit, !!!epsi))

data$lb = exp(data$alpha0 + data$beta0 + data$gamma0 +
              (data$alpha1 + data$gamma1) * data$tscaled +
              data$u + data$epsi)

data$W = sapply(c(1 : (I * J * K * T * N)),
                function(x) rpois(1, data$lb[x]))

dataYZ = data %>% group_by(ID, Block, Band, DPT) %>%
 summarise(Insecticide = unique(Insecticide),
           tscaled = unique(tscaled), st = unique(st), N = unique(N),
           Y = sum(W), Z = sum(W > 0)) %>% as.data.frame

dataW = data %>% select(- alpha0, - alpha1, - beta0, - gamma0, - gamma1,
                        - u, - epsi, - lb)

return(list("dataYZ" = dataYZ, "dataW" = dataW))
```

```
}
```

> **Note**
>
> The following code presents the inference for one seed. As the computing time is growing fast with the number of trials and the number of sampled values, we set here the number of trials to 10 (I = 10) and the number of adaptations and iterations to 2000.
> Our simulation results are obtained by running this code for I = 10, 20, 40, 80, for all seeds between 0 and 999, with a number of adaptations and iterations equal to 36000.
> Scenarios "50% Y - 0% Z" and "0% Y - 50% Z" are obtained from scenarios "100% Y - 0% Z" and "0% Y - 100% Z", respectively. For example "50% Y - 0% Z" scenario with 40 trials corresponds to "100% Y - 0% Z" with 20 trials.

```
J = 4;   T = 3;   N = 10;   K = 4;

Block = c(1 : K);   Band = c(1 : J);
Beet = c(1 : N);   DPT = seq(0, 12, length.out = T);

alpha0 = 0.5;   gamma0 = c(0, - 0.13, - 1.13, - 1.24);
alpha1 = 0.16;   gamma1 = c(0, 0.24, - 0.14, - 0.15);
sig0 = 1.87;   chi = 0.27;   eta = 0.98

##########################################################################

# Building scenarios ######################################################
seed = 1;   I = 10
data = suppressMessages(simu_data(seed = seed, I = I))

scenarios = list(
  Y = data$dataYZ %>% mutate(Z = NA),
  Z = data$dataYZ %>% mutate(Y = NA),
  YhalfZhalf = data$dataYZ %>% mutate(Y = ifelse(ID <= (I / 2), NA, Y),
                                      Z = ifelse(ID > (I / 2), NA, Z)),
  YhalfZ = data$dataYZ %>% mutate(Y = ifelse(ID <= (I / 2), NA, Y)),
  YZ = data$dataYZ,
  W = data$dataW
)
##########################################################################


# Inference ###############################################################
nadapt = 2000;   niter = 2000

res_inf = NULL

# Inference YZ ------------------------------------------------------------
ID = as.numeric(as.factor(as.character(scenarios$Y$ID)));
INSEC = as.numeric(as.factor(as.character(scenarios$Y$Insecticide)));
```

```
TIME = as.numeric(scenarios$Y$tscaled);
ST = as.numeric(as.factor(as.character(scenarios$Y$st)));

K = length(unique(ID));
L = length(unique(INSEC));
M = length(unique(ST))

TIME_unique = unique(scenarios$Y$tscaled)[2 : 3]
T = length(unique(TIME_unique))

for(i in (1 : 5)){
  Y = scenarios[[i]]$Y;   Z = scenarios[[i]]$Z;
  N = scenarios[[i]]$N;   Q = length(Y)

  data_jags = list(
    "Y" = Y, "Z" = Z, "Q" =  Q, "ID" = ID, "INSEC" = INSEC,
    "TIME" = TIME, "ST" = ST, "K" = K, "L" = L, "M" = M,
    "N" = N, "T" = T, "TIME_unique" = TIME_unique
  )

  model <- jags.model("jags_models/modelYZ.txt", data = data_jags,
                      n.chains = 2, n.adapt = nadapt)

  samples <- coda.samples(model,
                          variable.names = c("gamma0", "gamma1", "Eff"),
                          n.iter = niter, thin = 10)

  bind = list(samples);
  names(bind) = paste("samples", names(scenarios)[i], sep = "_")

  res_inf = res_inf %>% append(bind)
}

# Inference W ----------------------------------------------------------------
ID = as.numeric(as.factor(as.character(scenarios$W$ID)));
INSEC = as.numeric(as.factor(as.character(scenarios$W$Insecticide)));
TIME = as.numeric(scenarios$W$tscaled);
ST = as.numeric(as.factor(as.character(scenarios$W$st)));
SBIT = as.numeric(as.factor(as.character(scenarios$W$sbit)))

K = length(unique(ID));   L = length(unique(INSEC));
M = length(unique(ST));   X = length(unique(SBIT))

TIME_unique = unique(scenarios$W$tscaled)[2 : 3]
T = length(unique(TIME_unique))

W = scenarios$W$W;   Q = length(W)

data_jags = list(
  "W" = W, "Q" =  Q, "ID" = ID, "INSEC" = INSEC,
```

```
  "TIME" = TIME, "ST" = ST,  "SBIT" = SBIT, "K" = K, "L" = L,
  "M" = M, "X" = X, "T" = T, "TIME_unique" = TIME_unique
)

model <- jags.model("jags_models/modelW.txt", data = data_jags,
                    n.chains = 2, n.adapt = nadapt)

samples <- coda.samples(model,
                        variable.names = c("gamma0", "gamma1", "Eff"),
                        n.iter = niter, thin = 10)

res_inf = res_inf %>% append(list("samples_W" = samples))
##########################################################################


# Formatting results ######################################################
t = scenarios$Y$tscaled %>% unique
Eff_6_true = (1 - exp(gamma0[2 : J] + gamma1[2 : J] * t[2])) * 100
Eff_12_true = (1 - exp(gamma0[2 : J] + gamma1[2 : J] * t[3])) * 100
truth = c(Eff_6_true, Eff_12_true, gamma0[2 : J], gamma1[2 : J])

n_scenarios = length(scenarios)

esti = lapply(
  res_inf, function(x) summary(x)$statistics %>% as.data.frame %>%
    rownames_to_column %>%
    filter(!(grepl("gamma", rowname) & Mean == 0)) %>%
    select(Mean) %>% as.matrix %>% as.vector
)

b_inf = lapply(
  res_inf, function(x) summary(x)$quantiles %>% as.data.frame %>%
    rownames_to_column %>%
    filter(!(grepl("gamma", rowname) & `2.5%` == 0)) %>%
    select(`2.5%`) %>% as.matrix %>% as.vector
)

b_sup = lapply(
  res_inf, function(x) summary(x)$quantiles %>% as.data.frame %>%
    rownames_to_column %>%
    filter(!(grepl("gamma", rowname) & `2.5%` == 0)) %>%
    select(`97.5%`) %>% as.matrix %>% as.vector
)

parameters = summary(res_inf[[1]])$statistics %>% as.data.frame %>%
  rownames_to_column %>%
  filter(!(grepl("gamma", rowname) & Mean == 0)) %>%
  select(rowname) %>% as.matrix %>% as.vector

res = do.call(rbind,
```

```
                lapply(c(1 : n_scenarios),
                       function(x)
                         data.frame(Truth = truth, Scenario = names(scenarios)[x],
                                    I = I, seed = seed,
                                    value = esti[[x]], parameters = parameters,
                                    b_inf = b_inf[[x]], b_sup = b_sup[[x]]
                         )
                )
)
##########################################################################
```

3 Results

3.1 Results obtained with real data

We present here the results obtained with real data. First, we check the convergence of the model for the scenarios defined in Table 2. We then compare the estimated values and the credibility intervals of the treatment efficacy obtained in the different scenarios.

3.1.1 Model convergence and posterior distributions

Figure 2 presents the Markov chains associated with the model parameters and treatment efficacies for the different scenarios. The x-axis presents the iteration number and the y-axis presents the sampled value. Results show that the chains are well mixed. Figure 3 presents the Gelman-Rubin statistics associated with the model parameters and treatment efficacies as a function of the iterations, for the different scenarios. We observe that this statistics converges to 1, which indicates the convergence of the algorithm.

Table 4 gives a summary of the posterior distributions of the model parameters and treatment efficacies obtained for the "100% Y - 0% Z" scenario. The significantly positive value of α_1 indicates that the aphid numbers tend to increase with time in untreated plots. The relatively high value of σ_0 (posterior mean equal to 1.87) reveals a strong variability in aphid numbers between trials. The posterior mean value of χ (0.27) suggests that the treatment efficacy varies across trials. The γ_0 parameter is negative for all three treatments, indicating a negative effect of the treatments on the aphid numbers at the time of pesticide spray. The Movento and Teppeki treatments have a similar effect with a posterior mean for γ_0 equal to -1.13 and -1.24, and a standard deviation equal to 0.12 and 0.11, respectively. The effect of treatment Mavrik Jet is weaker as its posterior mean for γ_0 is equal to -0.13 and its standard deviation is equal to 0.16. The γ_1 posterior means are negative for Movento and Teppeki (-0.14 and -0.15), suggesting that the effect of these treatments tend to increase with time, but the posterior mean value is positive for Mavrik Jet (= 0.24), suggesting that the effect of this treatment may decrease with time. However, the 95% credibility intervals of γ_1 include zero and these parameters are not very accurately estimated.

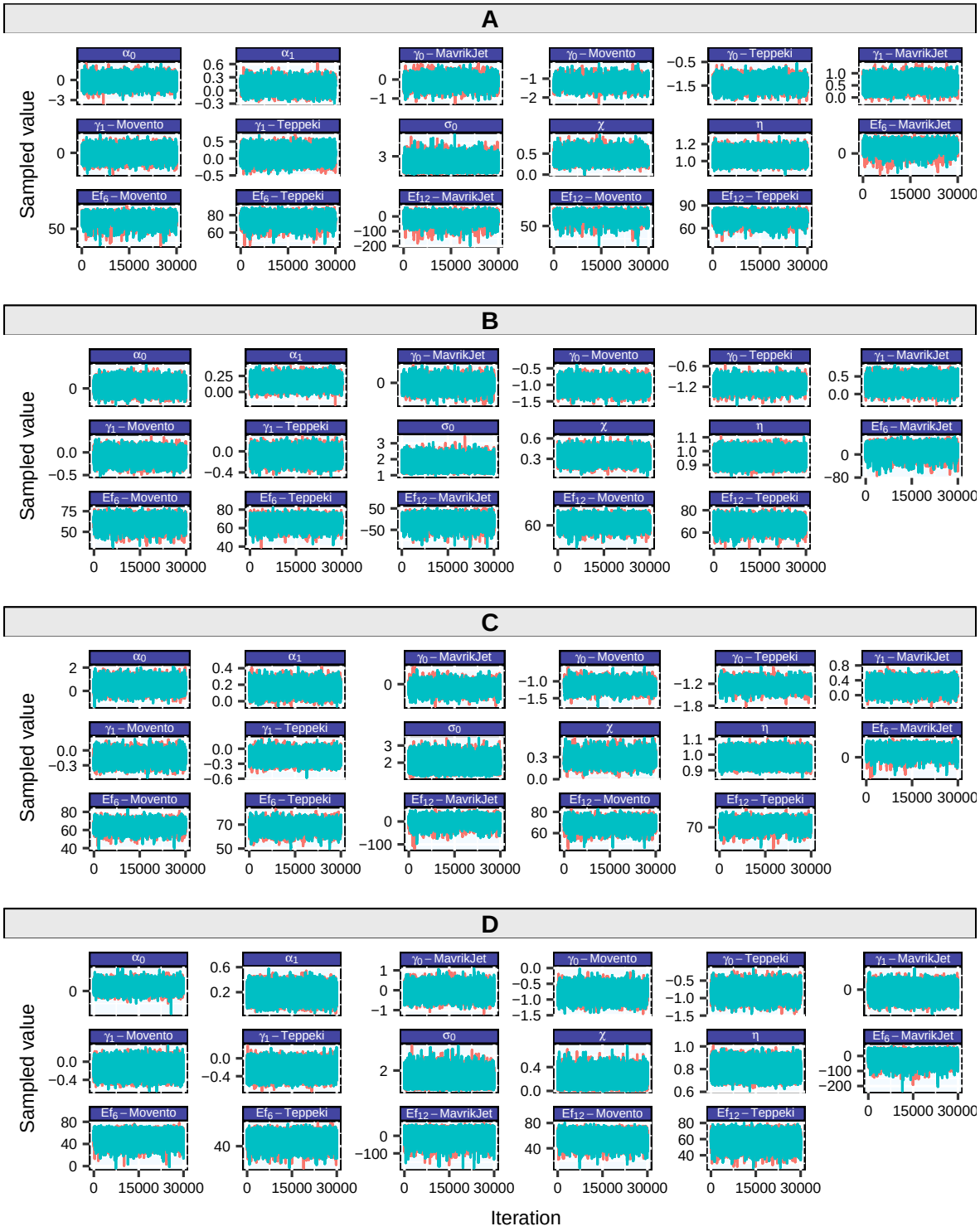

Figure 2: Model convergence - Markov chain for the model parameters and treatment efficacies, in the scenarios "50% Y - 0% Z" (**A**), "50% Y - 50% Z" (**B**), "100% Y - 0% Z" (**C**) and "0% Y - 50% Z" (**D**). The x-axis presents the iteration number and the y-axis presents the sampled value.

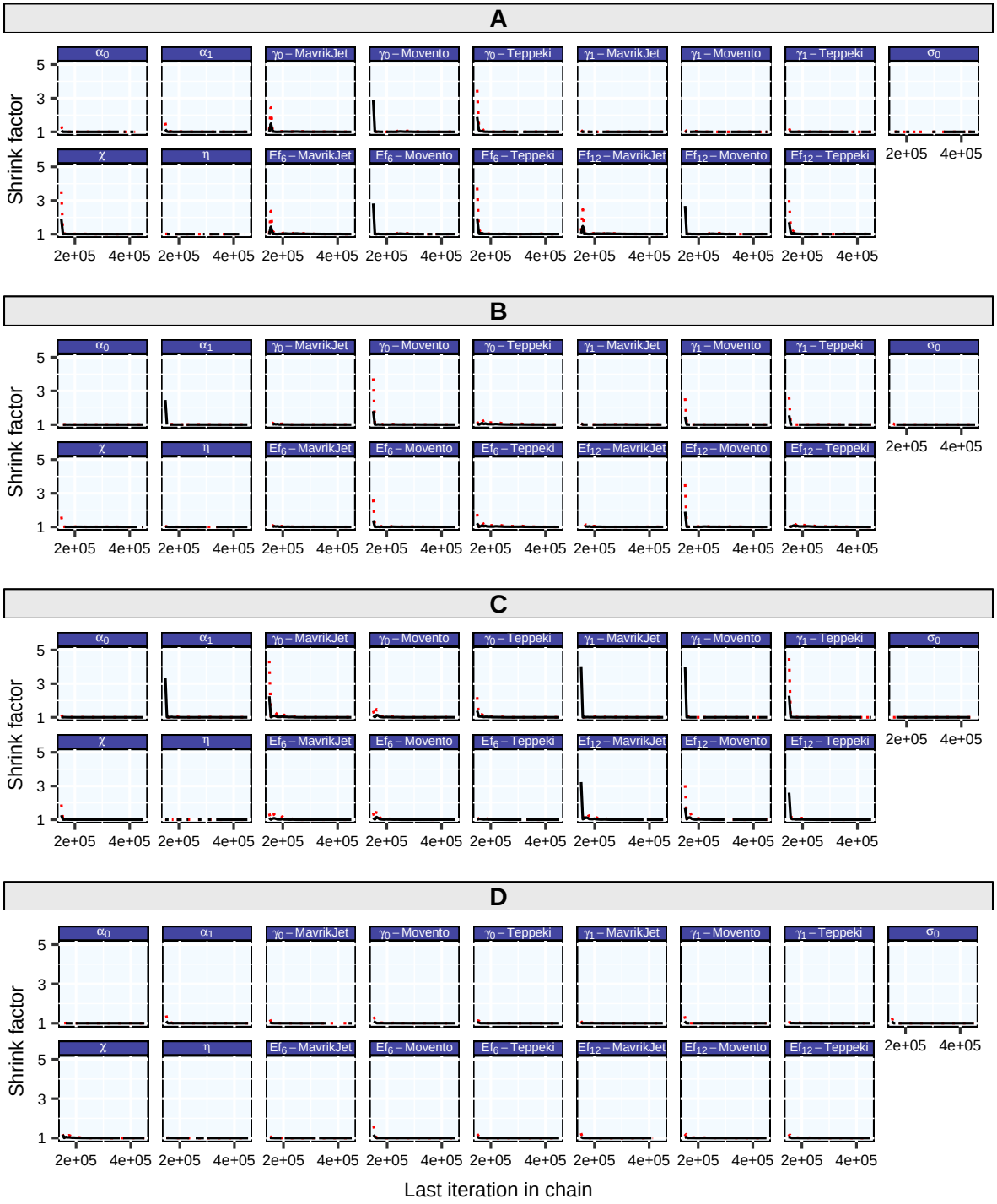

Figure 3: Model convergence - Gelman-Rubin statistics for the parameters of the model (Equation 1 - Equation 4) and for treatment efficacies (Equation 5), according to the scenarios "50% Y - 0% Z" (**A**), "50% Y - 50% Z" (**B**), "100% Y - 0% Z" (**C**) and "0% Y - 50% Z" (**D**). The x-axis presents the iteration number and the y-axis presents the Gelman-Rubin statistic.

Table 4: Summary of the posterior distributions obtained with the "100% Y - 0% Z" scenario for the model parameters and treatment efficacies: posterior mean, standard deviation, 2.5 and 97.5 quantiles.

Parameter	Mean	SD	2.5%	97.5%
α_0	0.49	0.34	-0.19	1.17
α_1	0.16	0.06	0.04	0.29
σ_0	1.87	0.26	1.44	2.46
χ	0.28	0.06	0.15	0.40
η	0.98	0.03	0.92	1.03
γ_0 - untreated	0.00	0.00	0.00	0.00
γ_0 - Mavrik Jet	-0.12	0.17	-0.46	0.20
γ_0 - Movento	-1.13	0.12	-1.37	-0.88
γ_0 - Teppeki	-1.24	0.11	-1.46	-1.02
γ_1 - untreated	0.00	0.00	0.00	0.00
γ_1 - Mavrik Jet	0.24	0.13	-0.01	0.49
γ_1 - Movento	-0.15	0.10	-0.33	0.04
γ_1 - Teppeki	-0.15	0.09	-0.32	0.02

3.1.2 Estimated values of pesticide treatment efficacies

Figure 4 presents the posterior means and the 95% credibility intervals of treatment efficacies at 6 days (A) and 12 days (B) after pesticide spray, for the "100% Y - 0% Z", "50% Y - 50% Z", "50% Y - 0% Z" and "0% Y - 50% Z" scenarios. Different scenarios are indicated by different colors. The x-axis presents the efficacy and the y-axis presents the treatments. Overall, the results obtained are consistent across scenarios; Teppeki and Movento show higher mean efficacies than Mavrik Jet, and the credibility intervals are narrower for Teppeki and Movento than for Mavrik Jet in all scenarios. The credibility interval of the "100% Y - 0% Z" scenario is narrower than that of the "50% Y - 50% Z" scenario, which is itself narrower than that of the "50% Y - 0% Z" and "0% Y - 50% Z" scenarios. Overall, the credibility interval sizes obtained with the "100% Y - 0% Z" scenario are 25% to 45% smaller than those obtained with the "50% Y - 0% Z" scenario (Table 5), aligning with the principle that increased data availability leads to more precise estimates. Results also indicate that credibility intervals are frequently larger with "0% Y - 50% Z" than with "50% Y - 0% Z," suggesting that more accurate estimates are achievable using Y compared to Z, at least in this specific case study. Interestingly, the sizes of the credibility intervals are approximately 25% smaller with "50% Y - 50% Z" compared to "50% Y - 0% Z," demonstrating that the combination of Y and Z observations collected from distinct trials proves beneficial and results in a reduction of uncertainty in the estimated treatment efficacy. This finding underscores the potential enhancement of treatment efficacy estimation through the combination of trials incorporating prevalence data and those incorporating intensity data.

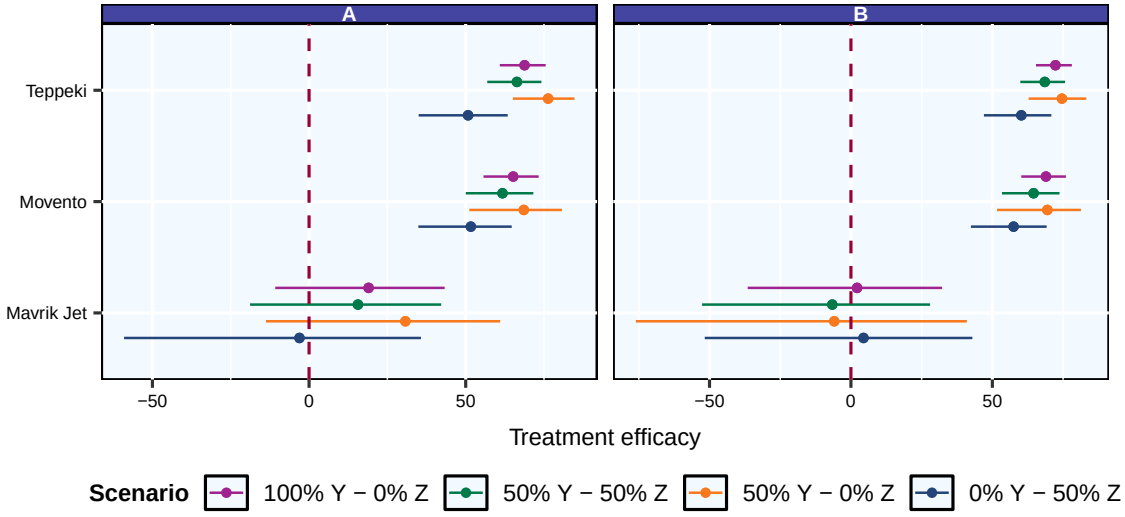

Figure 4: Estimated treatment efficacies after 6 days (**A**) and after 12 days (**B**), with their credibility intervals. Colors correspond to the different scenarios.

Table 5: Differences in the sizes of the 95 credibility intervals (CI) of the estimated treatment efficacies for the scenarios "50% Y - 50% Z", "0% Y - 50% Z" and "100% Y - 0% Z", compared to "50% Y - 0% Z". The difference is given in percentage. A positive (negative) value indicates an increase (decrease) of the credibility interval size. The third column indicates differences for the efficacy at 6 days after pesticide spray, and the fourth column indicates the difference for the efficacy at 12 days.

Insecticide	Data	For efficacy at 6 days	For efficacy at 12 days
Mavrik Jet	0% Y - 50% Z	26.7	-19.1
Movento	0% Y - 50% Z	0.8	-9.8
Teppeki	0% Y - 50% Z	44.4	16.9
Mavrik Jet	100% Y - 0% Z	-27.7	-41.3
Movento	100% Y - 0% Z	-40.5	-46.7
Teppeki	100% Y - 0% Z	-25.6	-37.8
Mavrik Jet	50% Y - 50% Z	-18.4	-31.1
Movento	50% Y - 50% Z	-27.1	-31.4
Teppeki	50% Y - 50% Z	-12.1	-22.4

3.2 Results obtained by simulation

3.2.1 Interest of combining trials with prevalence and trials with intensity

In this section, we consider the situation where only one type of observation is available per trial - pest prevalence or pest intensity. We compare the accuracy of the estimated parameters and estimated levels of treatment efficacy obtained by combining both types of trials compared to the results obtained using each set of trials separately.

The parameters used to generate the data are given in Table 6.

Table 6: Parameters used to generate virtual data

Parameters	α_0	α_1	γ_{00}	γ_{01}	γ_{02}	γ_{03}	γ_{10}	γ_{11}	γ_{12}	γ_{13}	σ_0	η	χ
Values	0.50	0.16	0.00	-0.13	-1.13	-1.24	0.00	0.24	-0.14	-0.15	1.87	0.98	0.27

Figure 5 represents the E_γ (Equation 6) (**A**), E_{Ef_6} (Equation 7) (**B**) and $E_{Ef_{12}}$ (Equation 7) (**C**) evaluation criteria for the "0% Y - 50% Z", "50% Y - 0% Z" and "50% Y - 50% Z" scenarios (different scenarios are indicated by different colors). The x-axis presents the number of trials and the y-axis the value of the criterion, averaged over the simulated data sets. For each number of trials and for each criterion, we observe that scenario "50% Y - 50% Z" gives a more accurate estimate than scenario "50% Y - 0% Z" which itself gives a more accurate estimate than scenario "0% Y - 50% Z". For example, for the efficacy at 6 days with 40 trials, the mean absolute error of scenario "50% Y - 0% Z" is 10% less than the mean absolute error of scenario "0% Y - 50% Z" (0.38 vs 0.42). The mean absolute error of scenario "50% Y - 50% Z" is 35% less than that of scenario "0% Y - 50% Z" (0.26 vs 0.42). The values of the three criteria decrease with the number of trials. The E_γ criterion decreases from 0.62 with 20 trials to 0.32 with 80 trials for the "50% Y - 50% Z" scenario. A value of 20 trials is therefore not sufficient to obtain an accurate estimate of the parameters.

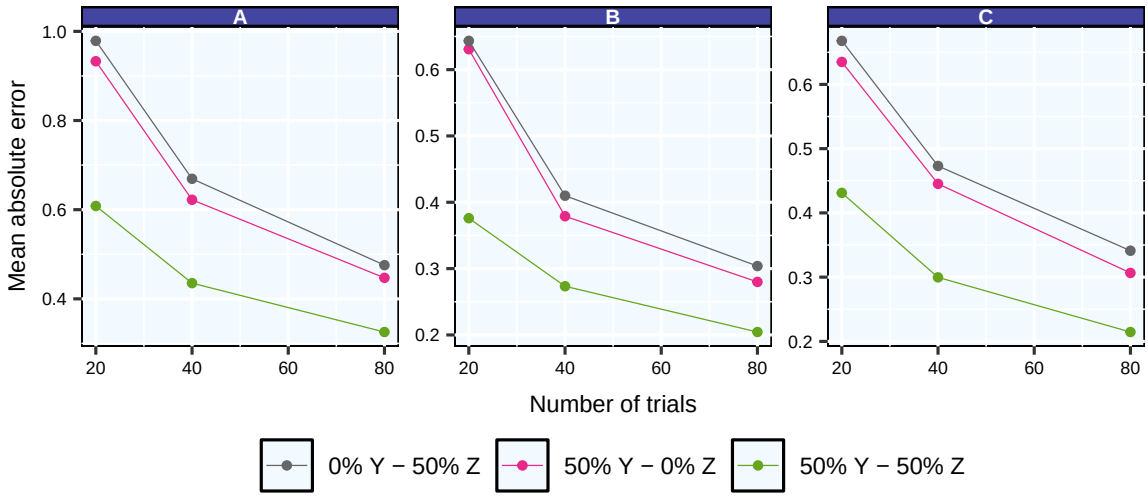

Figure 5: Values of the E_γ (Equation 6) (**A**), E_{Ef_6} (Equation 7) (**B**) and $E_{Ef_{12}}$ (Equation 7) (**C**) mean absolute error criteria for the "0% Y - 50% Z", "50% Y - 0% Z" and "50% Y - 50% Z" scenarios. The x-axis presents the number of trials and the y-axis presents the mean absolute error, averaged over the 974 simulated data sets. Different colors correspond to different scenarios.

Figure 6 presents the percentages of cases where the best treatment at 6 days (A) and 12 days (B) has been correctly identified for the "50% Y - 0% Z", "0% Y - 50% Z" and "50% Y - 50% Z" scenarios. The x-axis presents the number of trials and the y-axis the percentage of cases where the treatment identification is correct. In general, the best treatment is better identified when the number of trials increases. With the "50% Y - 50% Z" scenario, the best treatment at 6 days is well identified in 69% of cases with 20 trials and in 85% of cases with 80 trials. For each number of trials, the percentage of correctly identification is higher for the "50% Y - 50% Z" scenario than for the other two, and the scenario "50% Y - 0% Z" generally gives better results than the scenario "0% Y - 50% Z", except at 6 days with 20 trials. For example, at 12 days after treatment and with 40 trials, the percentage of correct identification is 5% higher with scenario "50% Y - 50% Z" than with scenario "50% Y - 0% Z" (78 vs 73), and 4% higher with the scenario "50% Y - 0% Z" than with scenario "0% Y - 50% Z" (73 vs

69). These results show the interest of combining prevalence and intensity data for assessing the efficacy of treatments and identifying the best treatments.

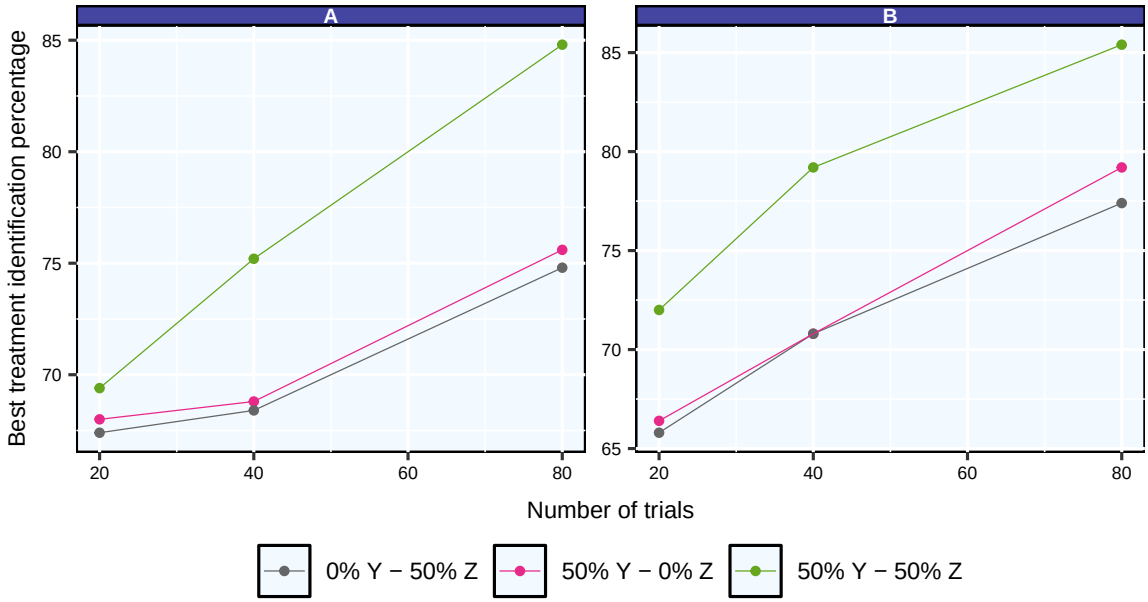

Figure 6: Comparison of proportion cases where the best treatment is correctly identified in the "0% Y - 50% Z", "50% Y - 0% Z" and "50% Y - 50% Z" scenarios. The x-axis represents the number of trials and the y-axis represents the percentage of cases where the best treatment has been correctly identified at 6 days (A) and 12 days (B), over the 974 simulated data sets. Different colors correspond to different scenarios.

3.2.2 Interest of adding intensity when prevalence is measured in all trials

We now consider a situation where prevalence is measured in each trial and intensity is measured in only some of these trials. We compare the results obtained when the data are combined and when they are used separately. As the prevalence data are usually more accessible in practice and the intensity data more costly, it is important to evaluate the interest of adding intensity data in the statistical analysis.

The parameters used to generate the data are the same as in 3.2.1.

Figure 7 presents the evaluation criteria E_γ (Fig 7A), E_{Ef_6} (Fig 7B) and $E_{Ef_{12}}$ (Fig. 7C) for the scenarios "0% Y - 100% Z", "50% Y - 0% Z", "50% Y - 100% Z" and "50% Y - 50% Z". The x-axis presents the number of trials and the y-axis presents the value of the criterion, averaged over the number of simulated data sets. Results show that the mean absolute errors are lower in scenarios "50% Y - 100% Z" and "50% Y - 50% Z" than in "0% Y - 100% Z", and that the mean absolute errors are lower in the "0% Y - 100% Z" scenario than in the "50% Y - 0% Z" scenario. For example, considering the treatment efficacy at 12 days with 40 trials (Fig. 7C), the mean absolute errors are 13% lower in scenarios "50% Y - 100% Z" and "50% Y - 50% Z" than in "0% Y - 100% Z" (0.30 vs 0.34), and the mean absolute error is 11% lower in "0% Y - 100% Z" than in "50% Y - 0% Z" (0.34 vs 0.38). Clearly, adding intensity data to prevalence data improves the accuracy of the estimations. The mean absolute errors decrease with the number of trials. For example, the E_γ criterion decreases from 0.62 with 20 trials to 0.32 with 80 trials for the "50% Y - 50% Z" scenario. As noted above, 20 trials is clearly not sufficient to obtain accurate results.

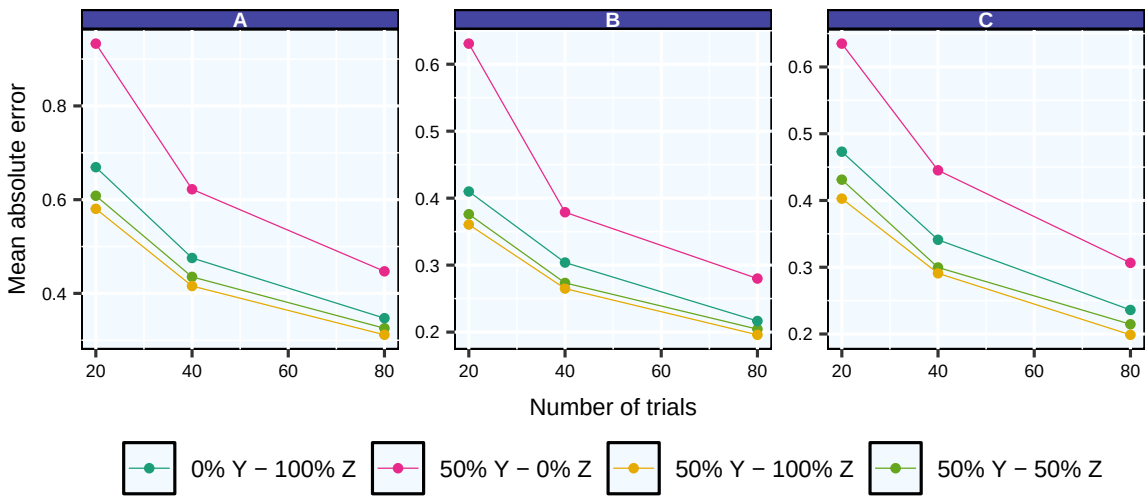

Figure 7: Values of E_γ (Equation 6) (**A**), E_{Ef_6} (Equation 7) (**B**) and $E_{Ef_{12}}$ (Equation 7) (**C**) for the "0% Y - 100% Z", "50% Y - 0% Z", "50% Y - 100% Z" and "50% Y - 50% Z" scenarios. The x-axis presents the number of trials and the y-axis presents the absolute error averaged over the 974 simulated data sets. Different colors correspond to different scenarios.

3.2.3 Is it better to measure intensity or prevalence in new pest surveys?

In order to optimize the design of new pest surveys that might be conducted in the future, we determine which type of observations should be favored. For that purpose, we compare the results obtained with the "100% Y - 0% Z", "0% Y - 100% Z", "100% Y / 100% Z" and "100% W" scenarios (recall that W represents the unobserved number of aphids on each plant in the sample (Equation 1)), for different values of α_0 that defines the average number of infested plants. With α_0 = -1, the proportion of infested plants is generally much lower than one, while with α_0 = 2, 100% of plants are generally infested. The case α_0 = 1 leads to intermediate levels of infestation.

The three parameter sets used to generate the data are given in Table 7 and are labeled A, B and C.

Table 7: Parameters considered for the design of future pest surveys."

Set	$\boldsymbol{\alpha_0}$	α_1	γ_{00}	γ_{01}	γ_{02}	γ_{03}	γ_{10}	γ_{11}	γ_{12}	γ_{13}	σ_0	η	χ
A	**-1.00**	0.16	0.00	-0.13	-1.13	-1.24	0.00	0.24	-0.14	-0.15	1.87	0.98	0.27
B	**1.00**	0.16	0.00	-0.13	-1.13	-1.24	0.00	0.24	-0.14	-0.15	1.87	0.98	0.27
C	**2.00**	0.16	0.00	-0.13	-1.13	-1.24	0.00	0.24	-0.14	-0.15	1.87	0.98	0.27

Figure 8 (A.1, B.1 and C.1) shows the mean absolute error E_γ (Equation 6) as a function of the number of trials for the four scenarios "100% Y - 0% Z", "0% Y - 100% Z", "100% Y / 100% Z" and "100% W". Figure 8 (A.2, B.2 and C.2) shows the distributions of infested plants with 40 trials corresponding to the three values of α_0 reported in Table 7. In case A (Table 7), the distribution of Z is such that Z is rarely close to 1 and often lower than 0.5 (Figure 8 A.2). In case C, the distribution of Z is such that Z is often very close to 1 (100% of plants infested). Case B is intermediate. The accuracy of the estimated values of the model parameters γ is better with scenario "100% Y - 0% Z" than with scenario "0% Y - 100% Z", for all number of trials. The advantage of "100% Y - 0% Z" is stronger in case of high pest prevalence (i.e., cases B and C) but very small in case of low pest prevalence (case A). For example, with 20 trials, the mean absolute error is 27% lower in the scenario "100% Y - 0%

Z" than in "0% Y - 100% Z" for parameter set C (0.55 vs. 0.75), 10% lower for parameter set B (0.55 vs. 0.62), and not different for parameter set A (0.64). The "100% W" scenario leads to similar results as "100% Y - 0% Z", regardless of α_0 and the number of trials. Results obtained with "100% Y / 100% Z" are generally similar to those obtained with "100% Y - 0% Z" and "100% W" but better than those obtained with the scenario "0% Y - 100% Z" in cases B and C. Here again, results show that 20 trials are not sufficient to obtain accurate parameter estimates.

4 Conclusion

In order to evaluate pest treatment efficacy, numerous trials are conducted to monitor pest prevalence and intensity. Quite often, only one type of data is available and, when both prevalence and intensity are available, they are usually analysed separately. In this paper, we propose an alternative approach based on a hierarchical statistical model able to analyze intensity and prevalence data, simultaneously.

We successfully apply the model to a real data set including prevalence and incidence data collected to evaluate three pesticide treatments against aphids in sugar beets. The model is fitted to this data set using a Markov chain Monte Carlo algorithm, and convergence was quickly achieved after a few thousands iterations. Results show that the use of both prevalence and intensity data led to a substantial reduction of the uncertainty in the parameter estimates, compared to the use of a single type of data.

Results obtained from simulated data confirm that, when pest prevalence and pest intensity are collected separately in different trials, the model parameters are more accurately estimated combining both prevalence and intensity trials than using one type of trials only. We also find that, when prevalence data are collected in all trials and intensity data are collected in a subset of trials, estimations and pest treatment ranking are more accurate using both types of data than using prevalence data only. Moreover, when only one type of observation can be collected in a pest survey or in an experimental trial, our analysis indicates that it is usually better to collect intensity data than prevalence data, especially in situations where all or most of the plants are expected to be infested. Finally, our simulations show that it is unlikely to obtain accurate results with fewer than 40 trials when assessing the efficacy of pest control treatments based on prevalence and intensity data.

Although our framework is illustrated to compare the efficacy of plant pest treatments, it could be applied to other areas of research in the future, in particular for optimizing designs used in animal and human epidemiology. It is imperative to note that the ultimate selection of a design should be contingent upon the consideration of local constraints. As the model codes are made fully available, we believe that these codes could be used by different institutes to compare many different designs in the future, not only the types of designs considered in our paper. Of particular significance is the capability of our model to optimize sample sizes, with its impact contingent on the relative importance of within-trial variability compared to between-trial variability.

5 Author contributions

AF and DM designed the study. AF performed the computations. AF and DM wrote the paper.

6 Funding

This work was partly funded by the project SEPIM (PNRI) and by the RMT SDMAA.

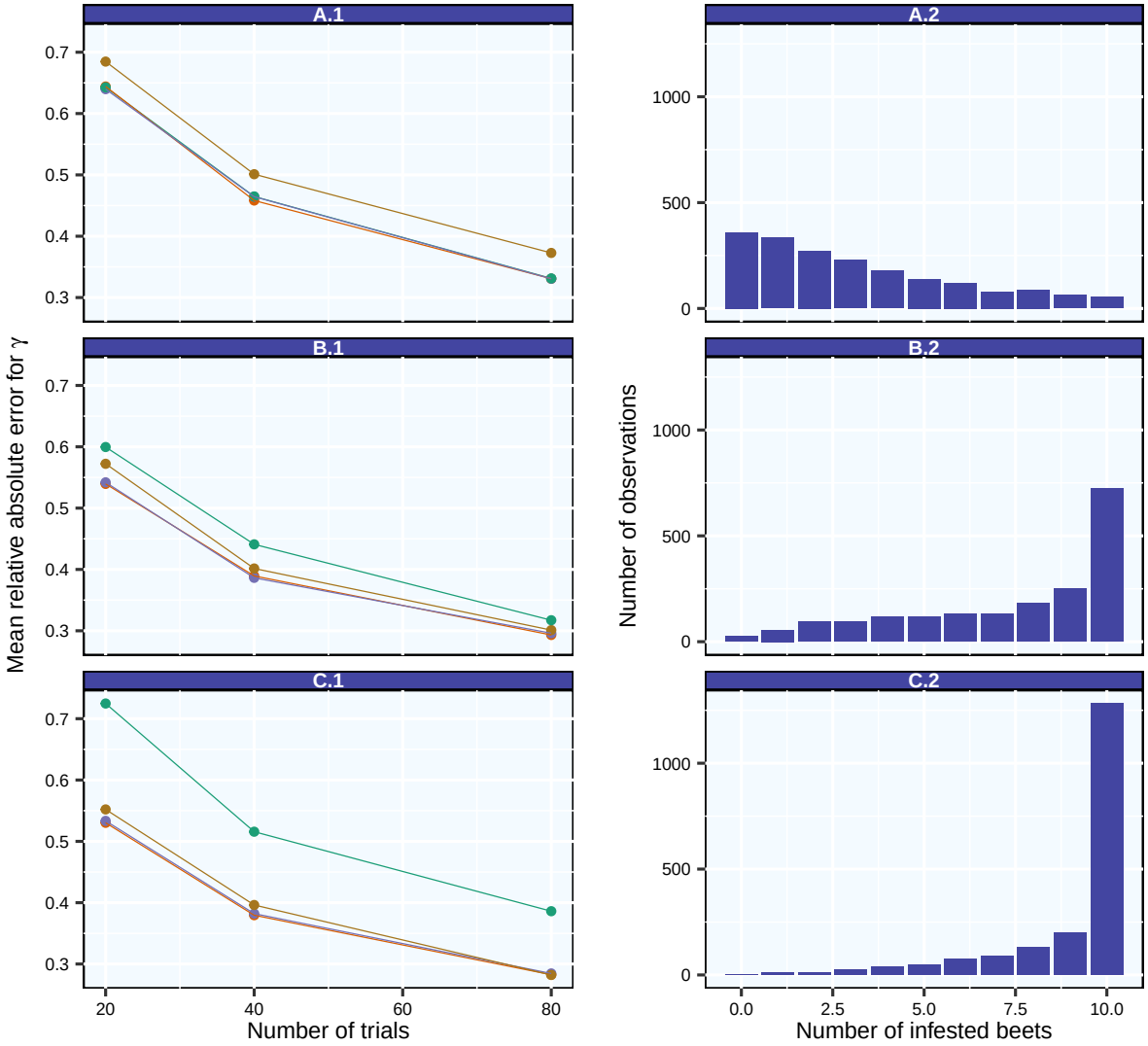

Figure 8: Comparison of the "100% Y - 0% Z", "0% Y - 100% Z", "100% Y / 100% Z" and "100% W" scenarios according to the distribution of Z and the number of trials, using the E_γ criterion (Equation 6). **A**, **B** and **C** correspond to different Z distributions which are given by A.2, B.2 and C.2 (distribution for a number of trials equal to 40). A, B and C respectively correspond to α_0 = -1, 1 and 2. The details of the simulation parameters are given in Table 7. A1, B1 and C1 represent the absolute error E_γ averaged over the 974 simulated data sets as a function of the number of trials. Colors correspond to the different scenarios.

7 Data availability

Simulated data and model parameters are available without restriction. The original experimental data may be available upon request.

8 Acknowledgements

We are grateful to Anabelle Laurent, Elma Raaijmakers, Kathleen Antoons and to the institute ITB (https://www.itbfr.org/) for their comments on this project.

We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing help and/or computing and/or storage resources.

The authors are thankful to the institutes that provided us with the data, namely the French Institut Technique de la Betterave, the sugar beet organisation of the Netherlands, and the Institut Royal Belge pour l'Amélioration de la Betterave.

Supplementary material

Model adjustment and comparison to the negative binomial model

To check the model's fit to the data, we performed a posterior predictive check of our model to check that the data were compatible with the model assumptions. To do so, we computed the probability of exceeding each individual data with the fitted model (2). Note that the number of pest individuals per plant are not available in practice; the data correspond to observed numbers of pest individuals for groups of N_i plants. Based on the posterior probability check (Figure 9), the computed probabilities were all falling in the range 0.22-0.93 (except for the observations equal to 0, for which the probability of being greater was equal to 1), and were thus not extreme. This result indicates that the model specified is not incompatible with the observed data and that the over-dispersion was correctly taken into account.

We also fitted a new model including a negative binomial distribution instead of a Poisson distribution. The results were almost identical between both types of model (Figure 10).

References

Agresti, Alan. 2015. *Foundations of Linear and Generalized Linear Models*. John Wiley & Sons.

Gelman, Andrew, John B Carlin, Hal S Stern, and Donald B Rubin. 1995. *Bayesian Data Analysis*. Chapman; Hall/CRC.

Gelman, Andrew, and Donald B Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science*, 457–72.

Gurmu, Shiferaw, and John Elder. 2000. "Generalized Bivariate Count Data Regression Models." *Economics Letters* 68 (1): 31–36.

Harrison, Xavier A. 2014. "Using Observation-Level Random Effects to Model Overdispersion in Count Data in Ecology and Evolution." *PeerJ* 2: e616.

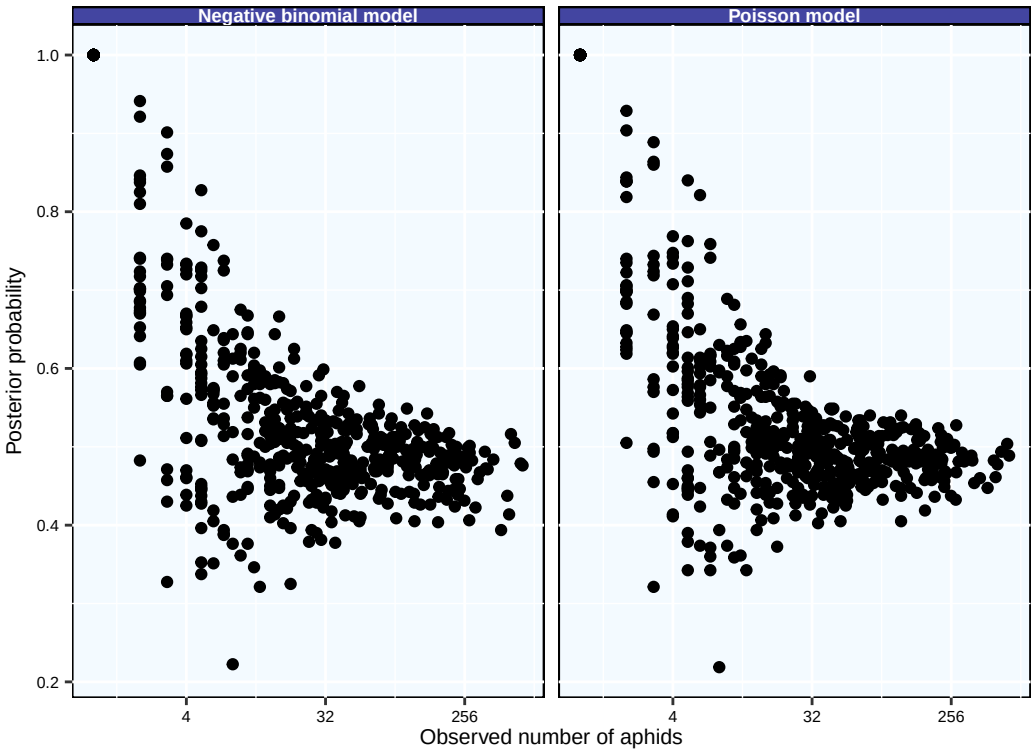

Figure 9: Posterior predictive check, The x-axis is on a logarithmic scale and represents the number of aphids increased by 1.

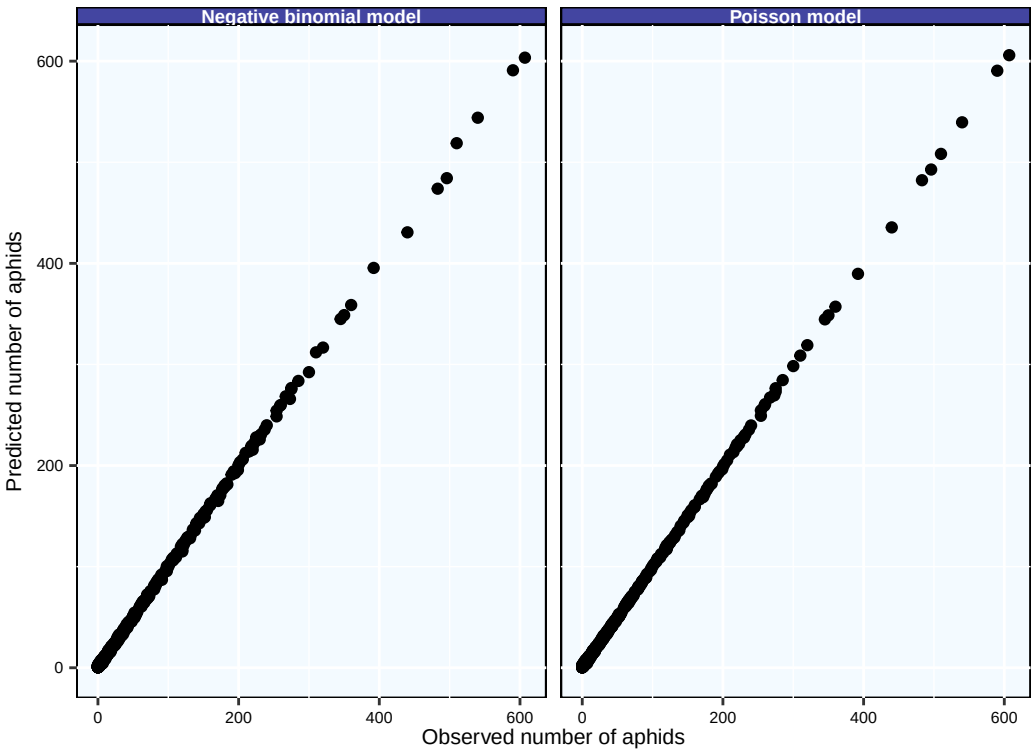

Figure 10: Observed vs predicted values.

Laurent, Anabelle, Armand Favrot, Fabienne Maupas, Cédric Royer, and David Makowski. 2023. "Assessment of Non-Neonicotinoid Treatments Against Aphids on Sugar Beets." *Crop Protection* 164: 106140. https://doi.org/10.1016/j.cropro.2022.106140.

Madden, LV, and G Hughes. 1999. "Sampling for Plant Disease Incidence." *Phytopathology* 89 (11): 1088–103.

Michel, Lucie, Francois Brun, and David Makowski. 2017. "A Framework Based on Generalised Linear Mixed Models for Analysing Pest and Disease Surveys." *Crop Protection* 94: 1–12.

Osiewalski, Jacek, and Jerzy Marzec. 2019. "Joint Modelling of Two Count Variables When One of Them Can Be Degenerate." *Computational Statistics* 34: 153–71.

Park, Eun Sug, and Dominique Lord. 2007. "Multivariate Poisson-Lognormal Models for Jointly Modeling Crash Frequency by Severity." *Transportation Research Record* 2019 (1): 1–6.

Pisa, Lennard W, Vanessa Amaral-Rogers, Luc P Belzunces, et al. 2015. "Effects of Neonicotinoids and Fipronil on Non-Target Invertebrates." *Environmental Science and Pollution Research* 22 (1): 68–102.

Plummer, Martyn. 2022. *Rjags: Bayesian Graphical Models Using MCMC*. https://CRAN.R-project.org/package=rjags.

Shaw, Allison K, Julie Sherman, F Keith Barker, and Marlene Zuk. 2018. "Metrics Matter: The Effect of Parasite Richness, Intensity and Prevalence on the Evolution of Host Migration." *Proceedings of the Royal Society B* 285 (1891): 20182147.

Wood, Thomas James, and Dave Goulson. 2017. "The Environmental Risks of Neonicotinoid Pesticides: A Review of the Evidence Post 2013." *Environmental Science and Pollution Research* 24 (21): 17285–325.