


Article

To Treat or Not to Treat Bees? Handy VarLoad: A Predictive Model for *Varroa destructor* Load

Hélène Dechatre ^{1,2}, Lucie Michel ³, Samuel Soubeyrand ³, Alban Maisonnasse ^{2,4}, Pierre Moreau ⁵, Yannick Poquet ^{1,2}, Maryline Pioz ^{1,2}, Cyril Vidau ^{2,6}, Benjamin Basso ^{1,2,6}, Fanny Mondet ^{1,2,*,†} 
and André Kretschmar ^{3,*,†}

- ¹ INRAE (Institut National de la Recherche Agronomique et de l'Environnement), Abeilles et Environnement, 84914 Avignon, France; helene.dechatre@gmail.com (H.D.); yannick.poquet@hotmail.com (Y.P.); maryline.pioz@inrae.fr (M.P.); benjamin.basso@inrae.fr (B.B.)
- ² UMT PrADE (Unité Mixte Technologique: Protection de l'Abeille dans l'Environnement), 84914 Avignon, France; a.maisonnasse.adapi@free.fr (A.M.); cyril.vidau@itsap.asso.fr (C.V.)
- ³ INRAE (Institut National de la Recherche Agronomique et de l'Environnement), BioSP, 84914 Avignon, France; lucie.michel@inrae.fr (L.M.); samuel.soubeyrand@inrae.fr (S.S.)
- ⁴ ADAPI (Association Pour le Développement de l'Apiculture en Provence), 13626 Aix en Provence, France
- ⁵ Beekeeper, 07270 Empurany, France; pierre-pm.moreau@laposte.net
- ⁶ ITSAP (Institut Technique et Scientifique de l'Abeille et de la Pollinisation), Institut de l'Abeille, 84914 Avignon, France
- * Correspondence: fanny.mondet@inrae.fr (F.M.); andre.kretschmar@inrae.fr (A.K.); Tel.: +33-432-722-175 (A.K.)
- † Equal contribution.



Citation: Dechatre, H.; Michel, L.; Soubeyrand, S.; Maisonnasse, A.; Moreau, P.; Poquet, Y.; Pioz, M.; Vidau, C.; Basso, B.; Mondet, F.; et al. To Treat or Not to Treat Bees? Handy VarLoad: A Predictive Model for *Varroa destructor* Load. *Pathogens* **2021**, *10*, 678. <https://doi.org/10.3390/pathogens10060678>

Academic Editor: Giovanni Cilia

Received: 21 April 2021

Accepted: 26 May 2021

Published: 30 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The parasitic *Varroa destructor* is considered a major pathogenic threat to honey bees and to beekeeping. Without regular treatment against this mite, honey bee colonies can collapse within a 2–3-year period in temperate climates. Beyond this dramatic scenario, *Varroa* induces reductions in colony performance, which can have significant economic impacts for beekeepers. Unfortunately, until now, it has not been possible to predict the summer *Varroa* population size from its initial load in early spring. Here, we present models that use the *Varroa* load observed in the spring to predict the *Varroa* load one or three months later by using easily and quickly measurable data: phoretic *Varroa* load and capped brood cell numbers. Built on 1030 commercial colonies located in three regions in the south of France and sampled over a three-year period, these predictive models are tools designed to help professional beekeepers' decision making regarding treatments against *Varroa*. Using these models, beekeepers will either be able to evaluate the risks and benefits of treating against *Varroa* or to anticipate the reduction in colony performance due to the mite during the beekeeping season.

Keywords: *Apis mellifera*; *Varroa destructor*; treatment; predictive model; beekeeping; decision-making tool

1. Introduction

The parasite *Varroa destructor* is considered a major pathogenic threat to honey bees [1] and to beekeeping. This mite is an ectoparasite affecting both adult bees and broods. Female mites have two distinct stages: a phoretic stage on adult bees and a reproductive stage, which takes place inside a capped brood during bee metamorphosis. The *Varroa* threat is not new for the beekeeping community, but with colony importations and the commerce of bees, this threat continues to increase. Indeed, these circumstances favor the *Varroa* spread throughout territories and the world's apiaries. This threat is all the more important given that the parasite spread is rapid [2]. Thus, bees and beekeepers cannot adapt and respond efficiently; on the contrary, *Varroa*, with continuous exposure to miticide treatments, responds with mechanisms of resistance [3–5]. Consequently, the current challenge is to develop new methods to limit *Varroa* numbers inside colonies.

Without regular efficient treatment against this mite, honey bee colonies can collapse within a 2–3-year period in temperate climates. Varroa feeding on pupal hemolymph can induce a decrease in adult bee body weight and malformations as well as reducing their life spans, thus weakening their immune systems [6–8]. Thus, it seems logical that infested colonies are less productive and efficient than healthy colonies, which can have significant economic impacts for beekeepers [9,10]. Beyond a threshold of 3 phoretic Varroa mites per 100 bees, the decrease in performance is correlated with the Varroa load [10]. According to this study, a colony with more than 3 phoretic Varroa mites per 100 bees produces, on average, 2.65 kg less honey than a colony below this threshold. Unfortunately, until now, it has not been possible to predict, from the mite population size in the spring, the population load in the summer, despite studies by Arechavaleta-Velasco and Guzman-Novoa (2001), Harris et al. (2003), and Lodesani et al. (2002), confirmed significant correlations between the amount of brood and/or the fertility of the mites [11–13] and population growth [1].

Models of Varroa dynamics have been previously established but mainly carried theoretical descriptions and only allowed for the evaluation of the instantaneous Varroa load. Wilkinson and Smith’s model [14] was built from virtual colonies, and DeGrandi-Hoffman and Curry’s model [15] was based on the BEEPOP honey bee colony population dynamics model [16]; a BEEHAVE Varroa unit was developed by Becher et al. [17]. These models were based on parametric values available from previous studies [16,18–26]. As these models primarily work with mathematical extrapolation, instead of being data-derived, we assumed that the resulting parametric values could be revised. Additionally, in the twenty years since these models were published, Varroa biology may have coevolved with its host. The coevolution between Varroa and honey bees has been reported by Kurze et al. [27] and includes host resistance behaviors, which involve a decrease in the Varroa reproduction rate as well as perturbations in the biological cycle of the mite. Moreover, previous studies serving as the basis for model construction were based on honey bees with different European origins and on Africanized honey bees [19]. Honey bee origins affect Varroa reproduction [28] and, consequently, Varroa population sizes. To increase its predictability, here, we used a model based on empirical data.

The most important information for a beekeeper is not the Varroa load at the time of honey flow because most treatment compounds, even some labeled “natural” (e.g., formic acid or thymol), are banned or not recommended during honey flow [29]. The aim of this study was therefore to predict the Varroa load one or three months later, from its baseline level in early spring, to anticipate colony performance for honey flow, knowing that the reduced performance threshold is 3 phoretic Varroa mites per 100 bees. Aimed as a useful tool for beekeepers, the model Handy Varload is based on inexpensive, accessible, and quickly measurable data in the field.

2. Results

2.1. Variable Selection (for Variable Definitions, See Materials and Methods, Statistical Analysis)

The variable “phoretic Varroa” measured at $t = 0$ was continuous with 25% of zeros, 27% of the data in $[0, 1]$, and 48% of the data in $[1, 30]$. The zero-inflated beta distribution is similar to the beta distribution but allows zeros as response values in which the ν parameter models the probability of obtaining zero. The distribution features of the variable “phoretic Varroa” (Vp_t) require dividing by 100 in order to fit the data to the interval $[0, 1]$. We then modeled this new response variable by a zero-inflated beta distribution, with parameter variation depending on covariates. A first model selection was performed to choose the best variables to model μ (see AICc comparisons in Table 1; more details are provided in Supplementary Materials Table S1). At the end of this preliminary selection, two models including the apiary random factor were retained, one for the 1-month adjustment and the other for the 3-month adjustment, noted (*) and (**) in Table 1, respectively.

Table 1. Comparisons of the tested models investigating the influence of phoretic Varroa numbers (per 100 bees) at $t = 0$, capped brood cell numbers, varbrood, and date of predicted phoretic Varroa numbers as a function of the estimation length, using the AICc criterion. $N = 867$ for data adjustment at one month ($x = 1$) and $N = 93$ for data adjustment at three months ($x = 3$).

Model	Adjustment for $x = 1$	Adjustment for $x = 3$
	AICc	AICc
phoretic Varroa	−2477.5	−268.3
capped brood cells	−2320.6	−261.8
varbrood (**)	−2540.1	−296.9
date	−2413.0	−260.9
phoretic Varroa + capped brood cells	−2488.0	−271.1
phoretic Varroa + date	−2561.7	−266.3
phoretic Varroa + varbrood	−2538.2	−295.7
capped brood cells + date	−2412.1	−261.4
capped brood cells + varbrood	−2580.1	−297.7
date + varbrood	−2618.2	−294.7
phoretic Varroa + capped brood cells + date	−2564.9	−269.1
phoretic Varroa + capped brood cells + varbrood	−2582.4	−296.0
phoretic Varroa + date + varbrood	−2616.2	−293.8
capped brood cells + date + varbrood (*)	−2645.5	−295.5
phoretic Varroa + capped brood cells + varbrood + date	−2647.8	−293.9
(*) and (**) + apiary random effect	−2651.1	−316.9

The final models (A and B, see below) were obtained after a second variable selection based on AICc comparisons, using the modeled σ and ν added to the (*) and (**) preliminary models. The number of phoretic Varroa present at t was modeled by the following zero-inflated beta models (BEZI in “*gamlss*”):

$$\begin{aligned} Vp_t &\sim \text{BEZI}(\mu, \sigma) && \text{with } (1 - \nu) \text{ probability} \\ Vp_t &= 0 && \text{with } \nu \text{ probability} \end{aligned}$$

For data adjustment at one month: (A)

$$\begin{aligned} \mu &= \text{logit}^{-1}(\alpha_0 + \alpha_1 Vb_{t-x} + \alpha_2 Cp_{t-x} + \alpha_3 D_t + Ap) \\ \sigma &= \exp(\beta_0 + \beta_1 Vb_{t-x} + \beta_2 D_t + Ap) \\ \nu &= \text{logit}^{-1}(\gamma_0 + \gamma_1 Vb_{t-x} + \gamma_2 Cp_{t-x} + \gamma_3 D_t + Ap) \end{aligned}$$

For data adjustment at three months: (B)

$$\begin{aligned} \mu &= \text{logit}^{-1}(\alpha_0 + \alpha_1 Vb_{t-x} + Ap) \\ \sigma &= \exp(\beta_0 + Ap) \\ \nu &= \text{logit}^{-1}(\gamma_0 + \gamma_1 Vb_{t-x} + \gamma_2 Vp_{t-x}) \end{aligned}$$

where the α , β , and γ parameters are coefficients used to model μ , σ , and ν , respectively. As a consequence of this second variable selection, the final AICc was -3179.2 (A) and -343.5 (B) (see details in Tables S2 and S3). Varbrood, which was retained by model selection in all cases except for σ of model B, appeared as the most important explanatory variable.

2.2. Goodness of Fit and Prediction Evaluation

2.2.1. Parameter Uncertainty

For models A and B, the α , β , and γ parameters associated with μ , σ , and ν were estimated and their $CI_{95\%}$ s were computed (see Table 2). Based on the intercept, we can note that varbrood had the largest influence on the data adjustment for each parameter of model A. Moreover, the order of influence of model covariates was the same regardless of

the parameter: varbrood > date > capped brood cells. For ν of model B, phoretic Varroa had a larger influence than varbrood. The $CI_{95\%}$ range as positively correlated with covariate weights, i.e., the greater the weight, the larger the uncertainty.

Table 2. Estimated coefficient and 95% confidence interval ($CI_{95\%}$) of models A and B investigating the influence of varbrood, capped brood cells, phoretic Varroa, and date on the number of phoretic Varroa mites for mu, sigma, and nu parameters.

Model	Parameter	Covariate	Estimated Coefficient	Lower 95% CI	Upper 95% CI
A	Mu	Intercept	−5.830	−6.021	−5.640
		varbrood	0.025	0.021	0.028
		capped brood cells	0.002	0.001	0.003
		date	0.014	0.012	0.015
	Sigma	Intercept	6.579	6.233	6.925
		varbrood	−0.023	−0.030	−0.016
		date	−0.018	−0.021	−0.015
	Nu	Intercept	2.073	1.475	2.672
		varbrood	−0.063	−0.087	−0.039
		capped brood cells	−0.003	−0.006	−0.001
		date	−0.032	−0.039	−0.025
	B	Mu	Intercept	−3.982	−4.175
varbrood			0.023	0.019	0.027
Sigma		Intercept	4.460	4.140	4.779
		Intercept	−0.701	−1.468	0.065
Nu		varbrood	−0.077	−0.167	0.012
		phoretic Varroa	−3.786	−10.747	3.176

Moreover, the apiary effect depended on the horizon of prediction. Thus, the mean apiary effect was zero with varying estimated standard deviations depending on the data adjustment; at one month, the estimated standard deviation was 0.285, with a standard deviation of this estimate of 0.798, and at three months, the estimated standard deviation was 0.681, with a standard deviation of this estimate of 0.914 (see Supplementary Materials Tables S2 and S3).

2.2.2. Prediction Quality

The prediction quality can be evaluated using confidence intervals and error rates of models. Table 3 shows that for cross-validation, 97.6% ($N = 4999$) of sampled phoretic Varroa mites were in their $CI_{95\%}$ with model A and 97.3% ($N = 2328$) with model B. These coverage rates are heterogeneous with respect to Vp_t : they overestimate the targeted values (95%, 70%, or 50%) when $Vp_t \leq 3$, they are consistent when $3 < Vp_t \leq 10$, and they are significantly lower than the targeted values when $Vp_t > 10$, which roughly corresponds to only 5–10% of the hives. These results hold approximately for all tackled cases (cross-validation and training validation; models A and B).

Table 3. Coverage rates of confidence intervals (CI_{95%}, CI_{70%}, and CI_{50%}) of Vp_t for both approaches, cross-validation and training validation, for models A and B. The coverage rate provides the proportion of times that the CI contains the true value of Vp_t . For each method and each model, numbers of observed hives are reported for each class of Vp_t .

Cross-Validation								
Model A	Model B	Observed Vp_t	Model A			Model B		
Observed Colony Numbers	Observed Colony Numbers		CI _{95%}	CI _{70%}	CI _{50%}	CI _{95%}	CI _{70%}	CI _{50%}
4999	2328	all	97.6	83.6	67.7	97.3	83.1	67.8
4027	1700	≤3	99.6	91.5	76.3	99.7	97.9	87.9
724	526	>3 and ≤10	92.7	53.7	34.5	99.8	51.1	16
248	102	>10	80.6	42.3	24.2	45.1	2	0
Training Validation								
Model A	Model B	Observed Vp_t	Model A			Model B		
Observed Colony Numbers	Observed Colony Numbers		CI _{95%}	CI _{70%}	CI _{50%}	CI _{95%}	CI _{70%}	CI _{50%}
1438	749	all	92.6	75.3	61.8	57.8	39	26
1140	546	≤3	95.9	82.6	69	61.2	44.1	29.9
229	137	>3 and ≤10	82.1	49.8	37.6	60.6	29.9	17.5
69	66	>10	72.5	39.1	21.7	24.2	15.2	12.1

Predicted quantiles were used as an indicator of the accuracy of the prediction aimed by the model, i.e., the proportion of hives to be treated against Varroa. Predicting values by simulation may be seen as minimizing the risk of an incorrect prediction (the risk of unnecessarily increasing the number of hives to be treated) or may be necessary to more accurately target the correctly predicted value (the risk of ignoring a proportion of hives which should be treated and which will not be). For model B, outputs are based on the average Varroa load in April of 0.7 phoretic Varroa mites per 100 bees [30] (quoted Vp_{t-x}) and the threshold of 3 phoretic Varroa mites per 100 bees at the beginning of summer [10] (quoted Vp_t). The model indicates for each colony whether or not to treat (prediction that the threshold will exceed three Varroa mites). Figure 1 describes two extreme situations that correspond to two treatment strategies. The first two strategies, represented by Q97.5 and Q85, are no-risk situations because the model indicated that all colonies are to be treated, and thus no risks are taken of having a colony that exceeds the threshold of three. In these cases, the input costs are great, and 73% of colonies are unnecessarily treated. The second strategy (Q50) is an attempt to justify no treatment, and it estimates the respective risk; it provides reasons not to treat 71% of colonies at the risk of not treating the 24% of colonies that need treatment. This could be seen as the price to pay for engaging in a process of decreasing inputs. Intermediate quantiles allow beekeepers to find correct indicators based on calculated trade-offs. For example, considering indicators for Q72 (or Q71.5), 27% of colonies observed exceed the threshold of three; the model predicted to treat 11% when necessary (10% for Q71.5) and 17% when not necessary (16% for Q71.5). In these cases, there were as many colonies that were treated when not necessary (17%—in orange) as colonies untreated when necessary (16%—in red) for Q72, and the inverse occurred for Q71.5 (Figure 1).

The first and third cases are the hives that are necessary to treat. The percentages of these four categories are provided for each level of risk.

This figure is based partly on Table S4 of Supplementary Materials; Tables S4 and S5 show all results for models A and B of the two model evaluations (cross-validation and training validation). For both models, the smaller the quantile, the lower the global error rate. For larger quantiles (Q97.5 and Q85), models predicted better Vp_t when the phoretic Varroa number exceeded the threshold of three Varroa mites at t . Model predictions of Vp_t were relatively good when the earlier phoretic Varroa number was at three, the maximum. However, models failed to produce correct predictions when the mite number at $t-x$ was higher than three for model A and higher than 0.7 for model B.

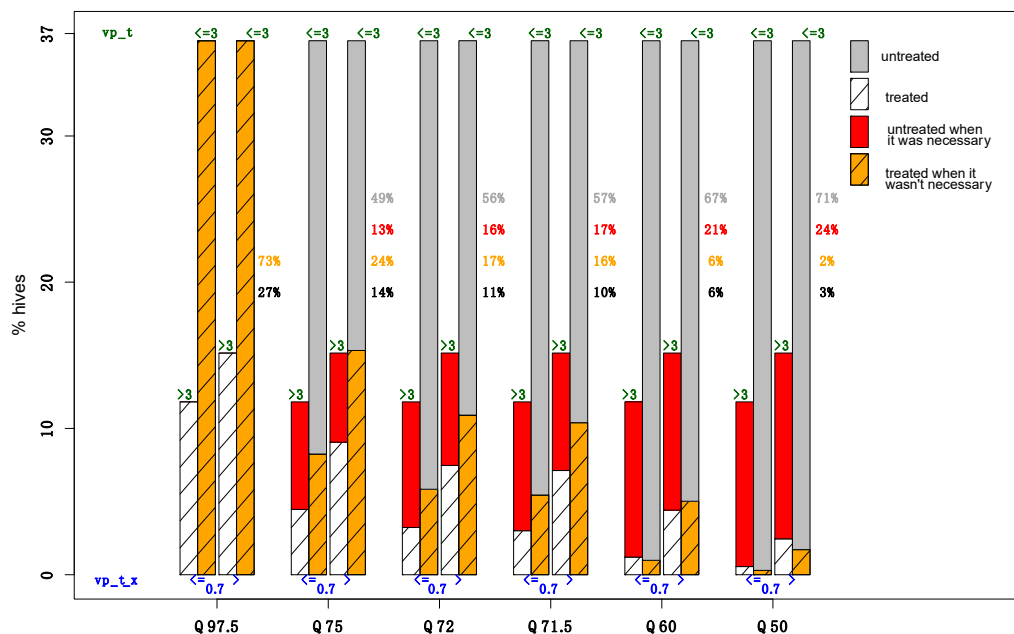


Figure 1. In this figure, 5 scenarios are presented with increasing risk (from left to right) taken by the beekeeper to not treat when the model predicts it was necessary or to treat when it was unnecessary. The risk is inversely proportional to the measure of quantile Q . For each level of risk, four cases are represented: (1) Hives with vp_t_x (i.e., Vp at $t = 0$) ≤ 0.7 and vp_t (i.e., Vp three months later) > 3 ; (2) Hives with vp_t_x (i.e., Vp at $t = 0$) ≤ 0.7 and vp_t (i.e., Vp three months later) ≤ 3 ; (3) Hives with vp_t_x (i.e., Vp at $t = 0$) > 0.7 and vp_t (i.e., Vp three months later) ≤ 3 ; (4) Hives with vp_t_x (i.e., Vp at $t = 0$) > 0.7 and vp_t (i.e., Vp three months later) > 3 .

3. Discussion

3.1. Selected Variables

The Handy VarLoad (HVL) model allowed for the prediction of the Varroa load at a given moment t , as a function of the previously observed Varroa load and of the available area for their reproduction, i.e., the number of honey bee brood cells.

Seasons influence the Varroa load, but only in the short term. This could be explained by the fact that, in one month, a beekeeper management intervention or a particular climatic event can have an effect on one or two Varroa generations, as the generation interval of capped brood is 12 days. The mite population growth rate is exponential during short periods (three months) and when mite populations are low; in contrast, Varroa population growth follows a logistic dynamic over longer periods (covering the entire production period) when density-dependent factors influence population growth [12]. Consequently, an event which increases or decreases Varroa reproduction may change the short-term Varroa load but have an insignificant influence on the long-term Varroa load. For example, disruption of honey bee colony broods could be offset by the Varroa population growth itself. Conversely, if colony brood disruption speeds up Varroa reproduction, the mite population size eventually stabilizes due to density dependence [25,31]. Moreover, during a three-month period, colonies undergo a series of favorable and unfavorable disruptions for Varroa development, particularly climatic, which will balance each other out. Finally, the apiary effect acts regardless of the delay between two phoretic Varroa measurements. Thus, the biological variability between colonies, the differences in management strategy between beekeepers, year, and region (climate) influence the Varroa load of the colony [32–35].

Contrary to previous mathematical models on Varroa load, the HVL model allows one to obtain a prediction with a measure of uncertainty, as well as the associated uncertainty for each parameter. The model uncertainty includes variability at the inter-apiary scale, in beekeeping management strategies, and in year and region effects. The apiary effect included in the model induces a large amount of prediction uncertainty, but, at the same

time, it assimilates the sampling diversity related to apiary characteristics (management, year, and region).

3.2. Beekeepers' Interest

The model presented here allows one to have a representation of the risk beekeepers take by not treating the apiary, according to the percentage of colonies that exceed the threshold of three Varroa mites. Different quantiles propose different decision-making indicators for beekeepers taking into account trade-offs between cost, time, and environmental effects of treatments, on the one hand, and the risk of losing infested colonies, on the other.

Moreover, beside economic trade-offs, Varroa treatments are not without consequences and, indeed, may induce acaricide resistance in Varroa [3–5], which is why beekeepers should treat only when economic risks are real. It is worth noting that treatments during the beekeeping season are not efficient over the long term [36]. These types of treatment must be used only when the aim is to temporarily decrease the Varroa load to optimize honey flow performance. Thus, this model takes into account integrated pest management.

The model can also be used to determine which apiaries should be given priority on lavender and sunflower fields if the spot number is limited. However, despite the fact that managing colonies at the apiary scale is more efficient, as honey bee colony performances are highly dependent on the characteristics of any apiary (Kretschmar et al., unpublished data), beekeepers may want to manage Varroa at the colony scale and thus strictly follow the model prediction.

3.3. Limits and Prospects of the Model

The choice to use only easy-to-measure variables in the field impairs the model's goodness of fit and, consequently, the estimation/prediction accuracy. Taking into account other variables (Varroa foundress density, Varroa infestation rate in the capped brood, natural death of Varroa mites measured on sticky boards, etc.) would have allowed better predicting the Varroa load. Including these additional variables in the present model could have easily improved its prediction power. Nevertheless, it would be far too long and complex to collect that type of data in the actual schedule of a beekeeper. If the sampling plan is unrealistic and impracticable at a large scale, the HVL model will be worthless. However, such improved models could be developed for researchers or technicians who work on a smaller scale and need to have better precision in their experimental frameworks. Another limit of this study is the sampled colony number: the more hives sampled, the better the estimation. In the present study, as the number of repetitions for each factor (management, year, and region) is limited, our sampling variation increased model uncertainty. Nevertheless, the Handy VarLoad model will be improved by the accumulation of data issued from the numerous experiments in which the two handy variables it uses (phoretic Varroa load and capped brood area) are commonly collected. As the database on which the model is based increases, the effect of covariates (apiary, region, season, beekeeping practices, etc.) can be better integrated.

4. Materials and Methods

4.1. Data Sampling

Data were collected from 310 colonies from 2014 to 2016 in three regions of France (PACA, AURA, and Occitanie; "dataset1") and from 720 colonies in 2018 in three regions of France (PACA, Nouvelle Aquitaine, and Centre; "dataset2"). Most of the colonies were kept on 10-frame Dadant hives and contained hybrid *Apis mellifera* L. queens. Colonies belonged to commercial beekeepers and thus displayed different sizes, dynamics, and management styles, which allowed us to take into account the variability which exists between beekeepers and apiaries. No treatment against Varroa was applied during the sampling periods.

At each sampling point, the amount of capped brood (noted *Cb*) was determined according to the ColEval method [37], and the phoretic mite load was estimated by sam-

pling around three hundred bees (or 45 g) from a frame containing an uncapped brood. Sampled bees were washed with a detergent solution and the number of Varroa mites retrieved (noted Vp) was counted [38]. Finally, to take into account seasonality, a “date” variable (noted D) was also created in which days were reported on a perpetual calendar with day 1 starting on 15 March of each year. This variable described the number of days ran from an initial time, which corresponds to the beginning of the measurable increase in the Varroa population after wintering. In our case, it corresponded approximately to the middle of March.

Sampling points were repeated at 30-day intervals, except for apiaries R16 to R18 (“dataset1”), in which measurements were sometimes performed every 12 days to mimic the generation interval of capped broods.

4.2. Statistical Methods

4.2.1. Distribution Adjustment on “dataset1”

All statistics were performed using the statistical software R version 3.3.0 [39]. Estimation of model parameters was carried out using the “gamlss” function of the eponymous package (Rigby and Stasinopoulos, 2005). The response variable (number of observed Varroa mites per 100 bees) was modeled with a generalized additive model for location, scale, and shape (GAMLSS). GAMLSS is an extension of the generalized linear model and the generalized additive model. It is a distribution-based approach to semiparametric regression models, in which all the parameters of the assumed distribution for the response can be modeled as additive functions of the explanatory variables, such as the *location* (e.g., *mean* μ), the *scale* (e.g., *variance* σ^2), the *shape* (skewness and kurtosis), and some *inflation* (e.g., at zero, ν). Moreover, we chose to use GAMLSS because it offers numerous choices for the distribution of the response variable and is suitable for time series data (Rigby and Stasinopoulos, 2001). GAMLSS was fitted to data using maximum (penalized) likelihood estimation implemented with the RS algorithm, which does not require accurate starting values for μ , σ , and ν to ensure convergence in comparison with the CG algorithm [40,41]. The most parsimonious model with the lowest corrected Akaike’s information criterion (AICc) [42], was selected; models with differences in AICc values lower than or equal to two were considered to be equivalent. We chose this selection criterion because, it is the most suitable criterion to model selection in predictive models for ecology and time series applications including forecasting [43]. Thus, it allows for the selection of the model that will best predict the response variable, i.e., the model with the best predictive accuracy.

Variables, which were described above, were transformed as follows to comply with the scaling conditions during model fitting:

$$Cb = \frac{Cb0}{100} \quad (1)$$

$$Vp = \frac{Vp0 * 100 * 0.14}{sw} \quad (2)$$

$$Vb = \log\left(\frac{Vp}{Cb + 130} * 100 + 1\right) * 50 \quad (3)$$

where Cb (Equation (1)) is a scaled value of the number of capped brood cells $Cb0$; Vp (Equation (2)) is the normalized rational number of Varroa mites for 100 honey bees (called “phoretic Varroa” in the present study), knowing that the weight per bee is 0.14 g, and sw in Equation (2) is the sampling weight of bees; Vb is a variable called “varbrood”, built to take into account the role of the amount of brood in the regulation of Varroa reproduction, and, more specifically, to integrate the fact that the more spread out the capped brood, the harder it is to capture phoretic Varroa mites hidden in the capped brood. The varbrood variable was thus obtained by taking the Neperian logarithm of the number of phoretic Varroa and dividing it by the number of capped brood cells. In Equation (3), 130 corresponds to the Cb median, 100 and 50 multipliers are necessary for the scale, and +1 is used to

avoid obtaining $\log(0)$. These three quantitative variables were mathematically reduced to the same scale, in order to be able to compare their respective weights during model adjustment. The date (measured as a number of days after the first measurement) was used without transformation.

The rational number of phoretic Varroa mites present at t (Vp_t) was modeled in the GAMLSS framework by a zero-inflated beta distribution with mean μ , standard deviation σ , and inflation at zero ν . Different specifications for μ , σ , and ν were used (see Results section). Our models were designed to predict Vp_t from explanatory variables typically collected at time $t-x$. Two horizons of prediction x were considered: a short-term horizon ($x = 1$ month, noted model A hereafter) and a long-term horizon ($x = 3$ months, noted model B hereafter). For $x = 1$ (model A), all data were used to fit the models (867 observations), whereas for $x = 3$ (model B), all the data providing this interval were used to avoid the use of time-overlapping pairs of observations (93 observations). Phoretic Varroa numbers, capped brood cell numbers, and varbrood present at $t-x$, as well as the date at t , were exploited as fixed factors; they are denoted by Vp_{t-x} , Cb_{t-x} , Vb_{t-x} , and D_t , respectively. Moreover, an «apiary» factor (noted Ap) was used as a random factor and includes the variability of the apiary, beekeeping management strategy, and year and region effects.

4.2.2. Goodness of Fit and Prediction Including “dataset2”

To assess the goodness of fit of the selected models, we explored the uncertainty of parameters and the prediction quality by comparing the predicted and observed values of Vp_t . We evaluated the prediction quality using two methods: cross-validation and training validation. In both methods, model performance was evaluated on data not included in the sample used to estimate model parameters.

For increasing the domain where the uncertainty of the parameters and the prediction quality of the models could be explored, a larger dataset (“dataset2”) was added to the first dataset (“dataset1”) with which the model parameters were estimated.

In the cross-validation method, observations of the hives of a given apiary were removed from the database, the model was fitted to the remaining data, and estimated parameters were plugged in to predict Vp_t for the hives of the apiary whose observations were removed (this case corresponds to predicting Vp_t for a new apiary based on observations collected from other apiaries). This procedure was repeated for each apiary of the dataset (i.e., 54 times for model A and 40 times for model B) and allowed us to provide averaged cross-validation assessments of the prediction performance.

The prediction performance was assessed with respect to the two following criteria:

- The actual coverage of 95%, 70%, and 50% confidence intervals of Vp_t (denoted by $CI_{95\%}$, $CI_{70\%}$, and $CI_{50\%}$), providing the proportion of times that the true value of Vp_t is contained within the CI;
- The use of different predicted quantiles of Vp_t (namely, $Q_{97.5\%}$, $Q_{85\%}$, $Q_{75\%}$, and $Q_{50\%}$) to evaluate the risk that the actual Vp_t exceeds the problematic threshold of 3 Varroa mites for 100 bees.

These criteria (CI and quantiles) were empirically calculated from 1000 simulations of the zero-inflated beta distribution in which the estimated values of μ , σ , and ν were inserted (random factors incorporated into μ , σ , and ν were randomly drawn at each simulation from centered normal distributions with standard deviations equal to their estimated values). Note that estimation uncertainty was neglected in this simulation procedure; this choice may lead to un-calibrated confidence intervals and quantiles. The comparisons of quantiles $Q_{97.5\%}$, $Q_{85\%}$, $Q_{75\%}$, and $Q_{50\%}$ with the threshold of 3 Varroa mites for 100 bees can be used as indicators to assess whether Vp_t will exceed this problematic threshold. The efficiency of these indicators was assessed with the error rate $\tau^{error}(\alpha)$ calculated as the ratio between (i) the number of hives for which the predicted quantile $Q_\alpha(Vp_t)$ is less than

or equal to 3 at time t , whereas the actual observation Vp_t is greater than 3, and (ii) the number of observations Vp_t greater than 3:

$$\tau^{error}(\alpha) = \frac{\sum_{R=1}^K \sum_{r=1}^{N_R} 1(Vp_{t,R,r} > 3, Q_{\alpha}^{-R}(Vp_{t,R,r}) \leq 3)}{\sum_{R=1}^K \sum_{r=1}^{N_R} 1(Vp_{t,R,r} > 3)}, \quad (4)$$

where K is the number of apiaries (54 for model A and 40 for model B), N_R is the number of hives in the apiary R (which ranges between 7 and 51), $Vp_{t,R,r}$ is the observed value of phoretic Varroa at time t for the hive r of the apiary R , and $Q_{\alpha}^{-R}(Vp_{t,R,r})$ is the predicted quantile at $\alpha\%$ of $Vp_{t,R,r}$ for the hives of the apiary R whose observations were removed ($-R$). The indicator function $E \rightarrow \mathbf{1I}$ takes the value of 1 if event E is true, or otherwise 0.

Equation (4) presents the case in which observations are greater than 3 and predictions are less than or equal to 3, and $\tau^{error}(\alpha)$ was also calculated when observations are less than or equal to 3 and predictions are greater than 3. The error rate $\tau^{error}(\alpha)$ can be computed in other specific conditions, for example, conditions related to the number of phoretic Varroa at $t-x$ ($Vp_{t-x,R,r}$).

In the training validation method, observations of the hives of all apiaries after a specific date, $t_A = 31$ (15 April) for model A, and $t_B = 118$ (11 July) for model B, were removed from the database, the model was fitted to remaining data, and estimated parameters were plugged in to predict Vp_t for the hives of all apiaries after t_A or t_B (this case corresponds to predicting Vp_t for an apiary already installed, based on observations collected beforehand from this apiary). This procedure was repeated for each year of the dataset (i.e., from 2014 to 2016 and 2018 for both models A and B) and allowed us to provide averaged training validation assessments of the prediction performance already introduced in the cross-validation approach.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/pathogens10060678/s1>, Table S1: Comparisons of the tested models investigating the influence of phoretic Varroa mite (per 100 bees) numbers, capped brood cell numbers, varbrood, and date on next phoretic Varroa numbers as a function of the estimation length, using the AICc criterion. $N = 867$ for data adjustment at one month ($x = 1$) and $N = 93$ for data adjustment at three months ($x = 3$), Table S2: R graphic output of model A summary, Table S3: R graphic output of model B summary, Table S4: Performance comparisons between different quantiles (Q97.5, Q85, Q75, Q50) for models A and B, depending on the observed phoretic Varroa numbers at $t-x$ and the observed phoretic Varroa numbers at t . Error rates represent the colony percentage for which the Varroa load at the horizon of prediction was badly predicted with the cross-validation method. For each quantile, the number of hives to treat depends on the error rate for which their percentages were reported, Table S5: Performance comparisons between different quantiles (Q97.5, Q85, Q75, Q50) for models A and B, depending on the observed phoretic Varroa numbers at $t-x$ and the observed phoretic Varroa numbers at t . Error rates represent the colony percentage for which the Varroa load at the horizon of prediction was badly predicted with the training validation method. For each quantile, the number of hives to treat depends on the error rate for which their percentages were reported.

Author Contributions: Conceptualization, F.M. and A.K.; methodology, F.M., A.K., L.M., H.D. and S.S.; investigation, P.M., A.M., C.V., B.B., F.M. and A.K.; data curation, H.D., L.M., S.S., Y.P., M.P. and A.K.; writing—original draft preparation, H.D., F.M. and A.K.; supervision, F.M. and A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available upon request to the corresponding author.

Acknowledgments: We wish to thank all beekeepers that gave us access to their apiaries to record data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rosenkranz, P.; Aumeier, P.; Ziegelmann, B. Biology and control of Varroa destructor. *J. Invertebr. Pathol.* **2010**, *103*, 96–119. [[CrossRef](#)]
2. Denmark, H.A.; Cromroy, H.L.; Cutts, L. *Varroa Mite. Varroa Jacobsoni Oudemans: (Acari: Varroidae)*; Fla. Department Agric. & Consumer Serv., Division of Plant Industry: Gainesville, FL, USA, 1991.
3. Baxter, J.; Eischen, F.; Pettis, J.; Wilson, W.T.; Shimanuki, H. Detection of fluvalinate-resistant Varroa mites in US honey bees. *Am. Bee J.* **1998**, *138*, 291.
4. Elzen, P.J.; Westervelt, D. Detection of coumaphos resistance in Varroa destructor in Florida. *Am. Bee J.* **2002**, *142*, 291–292.
5. Pettis, J.S. A scientific note on Varroa destructor resistance to coumaphos in the United States. *Apidologie* **2004**, *35*, 91–92. [[CrossRef](#)]
6. De Jong, D.; De Jong, P.H.; Goncalves, L.S. Weight loss and other damage to developing worker honeybees from infestation with *V. jacobsoni*. *J. Apic. Res.* **1982**, *21*, 165–216. [[CrossRef](#)]
7. Schneider, P.; Drescher, W. Einfluss der Parasitierung durch die Milbe *Varroa Jacobsoni* aus das Schlupfgewicht, die Gewichtsentwicklung, die Entwicklung der Hypopharynxdrüsen und die Lebensdauer von *Apis mellifera*. *Apidologie* **1987**, *18*, 101–106. [[CrossRef](#)]
8. Yang, X.; Cox-Foster, D.L. Impact of an ectoparasite on the immunity and pathology of an invertebrate: Evidence for host immunosuppression and viral amplification. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7470–7475. [[CrossRef](#)]
9. Emsen, B.; Guzman-Novoa, E.; Kelly, P.G. Honey production of honey bee (Hymenoptera: Apidae) colonies with high and low Varroa destructor (Acari: Varroidae) infestation rates in eastern Canada. *Can. Entomol.* **2014**, *146*, 236–240. [[CrossRef](#)]
10. Kretzschmar, A.; Maisonnasse, A.; Dussaubat, C.; Cousin, M.; Vidau, C. Performances des colonies vues par les observatoires de ruchers. *Innov. Agron.* **2016**, *53*, 81–93.
11. Arechavaleta-Velasco, M.E.; Guzman-Novoa, E. Relative effect of four characteristics that restrain the population growth of the mite *Varroa destructor* in honey bee (*Apis mellifera* L.) colonies. *Apidologie* **2001**, *32*, 157–174. [[CrossRef](#)]
12. Harris, J.W.; Harbo, J.R.; Villa, J.D.; Danka, R.G. Variable population growth of *Varroa destructor* (Mesostigmata: Varroidae) in colonies of honey bees hymenoptera: Apidae) during a 10-year period. *Environ. Entomol.* **2003**, *32*, 1305–1312. [[CrossRef](#)]
13. Lodesani, M.; Crailsheim, C.; Moritz, R.F.A. Effect of some characters on the population growth of mite *Varroa jacobsoni* in *Apis mellifera* L. colonies and results of a bi-directional selection. *J. Appl. Entomol.* **2002**, *126*, 130–137. [[CrossRef](#)]
14. Wilkinson, D.; Smith, G.C. A model of the mite parasite, *Varroa destructor*, on honeybees (*Apis mellifera*) to investigate parameters important to mite population growth. *Ecol. Model.* **2002**, *148*, 263–275. [[CrossRef](#)]
15. DeGrandi-Hoffman, G.; Curry, R. A mathematical model of Varroa mite (*Varroa destructor* Anderson and Trueman) and honeybee (*Apis mellifera* L.) population dynamics. *Int. J. Acarol.* **2004**, *30*, 259–274. [[CrossRef](#)]
16. DeGrandi-Hoffman, G.; Roth, S.A.; Loper, G.L.; Erickson, E.H., Jr. BEEPOP: A honeybee population dynamics simulation model. *Ecol. Model.* **1989**, *45*, 133–150. [[CrossRef](#)]
17. Becher, M.A.; Grimm, V.; Thorbek, P.; Horn, J.; Kennedy, P.J.; Osborne, J.L. BEEHAVE: A systems model of honeybee colony dynamics and foraging to explore multifactorial causes of colony failure. *J. Appl. Ecol.* **2014**, *51*, 470–482. [[CrossRef](#)] [[PubMed](#)]
18. Jay, S.C. The development of honeybees in their cells. *J. Apic. Res.* **1963**, *2*, 117–134. [[CrossRef](#)]
19. De Jong, D.; De Jong, P.H. Longevity of Africanized honey bees (Hymenoptera: Apidae) infested by *Varroa jacobsoni* (Parasitiformes: Varroidae). *J. Econ. Entomol.* **1983**, *76*, 766–768. [[CrossRef](#)]
20. Kovac, H.; Crailsheim, K. Lifespan of *Apis mellifera carnica* Pollm. infested by *Varroa jacobsoni* Oud. In relation to season and extent of infestation. *J. Apic. Res.* **1988**, *27*, 230–238. [[CrossRef](#)]
21. Fries, I.; Camazine, S.; Sneyd, J. Population dynamics of *Varroa jacobsoni*: A model and a review. *Bee World* **1994**, *75*, 4–28. [[CrossRef](#)]
22. Donzé, G.; Herrmann, M.; Bachofen, B.; Guerin, P.M. Effect of mating frequency and brood cell infestation rate on the reproductive success of the honeybee parasite *Varroa jacobsoni*. *Ecol. Entomol.* **1996**, *21*, 17–26. [[CrossRef](#)]
23. Martin, S.J.; Kemp, D. Average number of reproductive cycles performed by *Varroa jacobsoni* in honey bee (*Apis mellifera*) colonies. *J. Apic. Res.* **1997**, *36*, 113–123. [[CrossRef](#)]
24. Calis, J.N.M.; Fries, I.; Ryrle, S.C. Population modeling of *Varroa jacobsoni* Oud. *Apidologie* **1999**, *30*, 111–124. [[CrossRef](#)]
25. Martin, S.J. A population model for the ectoparasitic mite *Varroa jacobsoni* in honey bee (*Apis mellifera*) colonies. *Ecol. Model.* **1998**, *109*, 267–281. [[CrossRef](#)]
26. Martin, S.J. The role of Varroa and viral pathogens in the collapse of honey bee colonies: A modeling approach. *J. Appl. Entomol.* **2001**, *38*, 1082–1093.
27. Kurze, C.; Routtu, J.; Moritz, R.F. Parasite resistance and tolerance in honeybees at the individual and social level. *Zoology* **2016**, *119*, 290–297. [[CrossRef](#)]
28. Camazine, S. Differential reproduction of the mite *Varroa jacobsoni* (Mesostigmata: Varroidae), on Africanized and European honey bees (Hymenoptera: Apidae). *Ann. Entomol. Soc. Am.* **1986**, *79*, 801–803. [[CrossRef](#)]
29. Bogdavorov, S.; Charrière, J.-D.; Imdorf, A.; Kilchenmann, V.; Fluri, P. Determination of residues in honey after treatments with formic and oxalic acid under field conditions. *Apidologie* **2002**, *33*, 399–409. [[CrossRef](#)]

30. Maisonnasse, A.; Frontero, L.; Kretzschmar, A. Mesurer le taux de VP/100ab (Varroa phorétique pour 100 abeilles) dans les ruchers pour optimiser la gestion et la production. In Proceedings of the 5ème Journée de la Recherche Apicole, Paris, France, 8–9 February 2017.
31. Sumpter, D.J.T.; Broomhead, D.S. Relating individual behaviour to population dynamics. *Proc. R. Soc. Lond. B Biol. Sci.* **2001**, *268*, 925–932. [[CrossRef](#)]
32. Eguaras, M.; Marcangeli, J.; Oppedisano, M.; Fernández, N. Seasonal changes in *Varroa jacobsoni* reproduction in temperate climates of Argentina. *Bee Sci.* **1994**, *3*, 120–123.
33. Garcia-Fernandez, P.; Rodriguez, R.B.; Orantesbermejo, F.J. Influence of climate on the evolution of the population-dynamics of the *Varroa* mite on honeybees in the south of Spain. *Apidologie* **1995**, *26*, 371–380.
34. Kraus, B.; Velthuis, H.H.W. High humidity in the honey bee (*Apis mellifera* L.) brood nest limits reproduction of the parasitic mite *Varroa jacobsoni* Oud. *Naturwissenschaften* **1997**, *84*, 217–218. [[CrossRef](#)]
35. Currie, R.W.; Tahmasbi, G.H. The ability of high-and low-grooming lines of honey bees to remove the parasitic mite *Varroa destructor* is affected by environmental conditions. *Can. J. Zool.* **2008**, *86*, 1059–1067. [[CrossRef](#)]
36. Mondet, F.; Maisonnasse, A.; Kretzschmar, A.; Alaux, C.; Vallon, J.; Basso, B.; Dangleant, A.; Le Conte, Y. Varroa: Son impact, les méthodes d'évaluation de l'infestation et les moyens de lutte. *Innov. Agron.* **2016**, *53*, 63–80.
37. Hernandez, J.; Maisonnasse, A.; Cousin, M.; Beri, C.; Le Quintrec, C.; Bouetard, A.; Castex, D.; Decante, D.; Serval, E.; Buchwalder, G.; et al. ColEval: Honeybee COLony Structure EVALuation for Field Surveys. *Insects* **2020**, *11*, 41. [[CrossRef](#)] [[PubMed](#)]
38. Dietemann, V.; Nazzi, F.; Martin, S.J.; Anderson, D.L.; Locke, B.; Delaplane, K.S.; Wauquiez, Q.; Tannahill, C.; Frey, E.; Ziegelmann, B.; et al. *BEEBOOK Volume II. Standard Methods for Varroa Research 4.2.3.1.2.3*; University of Bern: Bern, Switzerland, 2012.
39. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: <https://www.R-project.org/> (accessed on 25 May 2021).
40. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape, (with discussion). *Appl. Stat.* **2005**, *54*, 507–554. [[CrossRef](#)]
41. Rigby, R.A.; Stasinopoulos, D.M. The GAMLSS project: A flexible approach to statistical modelling. In *New Trends in Statistical Modelling, Proceedings of the 16th International Workshop on Statistical Modelling, Odense, Denmark, 2–6 July 2001*; Klein, B., Korsholm, L., Eds.; Odense University Press: Odense, Denmark, 2001; pp. 249–256.
42. Hurvich, C.M.; Tsai, C.-L. Regression and Time Series Model Selection in Small Samples. *Biometrika* **1989**, *76*, 297–307. [[CrossRef](#)]
43. Aho, K.; Derryberry, D.; Peterson, T. Model selection for ecologists: The worldviews of AIC and BIC. *Ecology* **2014**, *95*, 631–636. [[CrossRef](#)]