Stoch Environ Res Risk Assess (2010) 24:455–462
DOI 10.1007/s00477-009-0333-z

ORIGINAL PAPER

Bayesian analysis of a dynamical model for the spread of the Usutu virus

Jenő Reiczigel · Katharina Brugger · Franz Rubel · Norbert Solymosi · Zsolt Lang

Published online: 13 August 2009

Abstract The Usutu virus is an arbovirus transmitted by mosquitoes and causing disease in birds. The virus was detected in Austria for the first time in 2001, while a major outbreak occurred in 2003. Rubel et al. (2008) developed a nine-compartment deterministic SEIR model to explain the spread of the disease. We extended this to a hierarchical Bayes model assuming random variation in temperature data, in reproduction data of birds, and in the number of birds found to be infected. The model was implemented in R, combined with the FORTRAN subroutine for the original deterministic model. Analysis was made by MCMC using a random walk Metropolis scheme. Posterior means, medians, and credible intervals were calculated for the parameters. The hierarchical Bayes approach proved to be fruitful in extending the deterministic model into a stochastic one. It allowed for Bayesian point and interval estimation and quantification of uncertainty of predictions. The analysis revealed that some model parameters were not identifiable; therefore we kept constant some of them and analyzed others conditional on them. Identifiability problems are common in models aiming to mirror the mechanism of the process, since parameters with natural interpretation are likely to exhibit interrelationships. This study illustrated that Bayesian modeling combined with conditional analysis may help in those cases. Its application to the Usutu model improved model fit and revealed the structure of interdependencies between model parameters: it demonstrated that determining some of them experimentally would enable estimation of the others, except one of them, from available data.

Keywords Dynamical model · Epidemiology · Hierarchical Bayes model · MCMC · Usutu virus

J. Reiczigel (✉) · Z. Lang
Department for Biomathematics and Informatics,
Faculty of Veterinary Science, Szent István University,
István u. 2, 1078 Budapest, Hungary
e-mail: reiczigel.jeno@aotk.szie.hu

J. Reiczigel · N. Solymosi
HAS-BCU Adaptation to Climate Change Research Group,
Villányi út 29-43, 1118 Budapest, Hungary

J. Reiczigel · K. Brugger · F. Rubel
Institute for Veterinary Public Health,
University for Veterinary Medicine Vienna,
Veterinärplatz 1, 1210 Vienna, Austria

1 Introduction

The Usutu virus (USUV) is an arbovirus transmitted by mosquitoes and causing disease in birds. USUV was distributed in the tropical regions of Africa until it was first detected in 2001 in Vienna, a region characterized by Cfb-climate (warm temperate, fully humid with warm summers) after Köppen (Kottek et al. 2006). A major outbreak occurred during the extraordinary hot summer in 2003. For details on its ecology and spread in central Europe see Weissenböck et al. (2002), Bakonyi et al. (2007) or Brugger and Rubel (2009). The histogram in Fig. 1 shows the numbers of USUV-positive dead blackbirds (*Turdus merula*) found in Eastern Austria from 2001 to 2005. Monitoring took place each year in July, August, and September (Chvala et al. 2007).

Rubel et al. (2008) developed a nine-compartment SEIR (susceptible-exposed-infected-removed) model to explain the spread of disease (Fig. 2), which was also used to investigate USUV dynamics for various climate change scenarios (Brugger and Rubel 2009). For an introduction to infectious disease models and a recent review of methods,

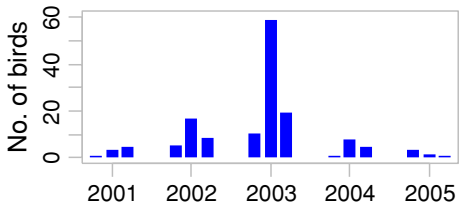

Fig. 1 Number of USUV-positive dead blackbirds (*Turdus merula*) found in Eastern Austria

see Keeling and Rohani (2007). The model in Fig. 2 has four compartments for mosquitoes (*L*-larvae, *S*-susceptible, *E*-exposed, that is infected but not yet infectious, *I*-infectious), and five for birds (*S*-susceptible, *E*-exposed, *I*-infectious, *R*-recovered (immune), *D*-dead). The model assumes climate-dependent logistic population growth and cross-infection between birds and mosquitoes. Naming conventions for the parameters are as follows: b and m denote birth and mortality rates, N and K denote population numbers and carrying capacity of the environment. Indices L, M, and B refer to parameters of larvae, mosquitoes, and birds, respectively. Temperature-dependence of a parameter is indicated by (T). For example, $b_L(T)$ denotes the birth rate of larvae, which is a temperature-dependent parameter. Concerning cross-infection, $k(T)$ denotes the temperature-dependent biting rate of mosquitoes, p_B and p_M stand for transmission probabilities, while δ_M refers to the proportion of non-diapausing (active) mosquitoes. Finally, $\gamma_M(T)$ and γ_B denote the transition rates from exposed to infectious, α_B stands for the removal rate to recovery or death, and v_B defines the proportion of the latter.

According to the model diagram in Fig. 2, there are four differential equations for mosquitoes.

$$\frac{dL_M}{dt} = (b_L(T)\delta_M N_M - m_L(T)L_M)(1 - L_M/K_M) - b_M(T)L_M \quad (1)$$

$$\frac{dS_M}{dt} = -k(T)p_B\,\delta_M I_B/K_B S_M + b_M(T)L_M - m_M(T)S_M \quad (2)$$

$$\frac{dE_M}{dt} = k(T)p_B\,\delta_M I_B/K_B S_M - \gamma_M(T)E_M - m_M(T)E_M \quad (3)$$

$$\frac{dI_M}{dt} = \gamma_M(T)E_M - m_M(T)I_M \quad (4)$$

and five equations for birds.

$$\frac{dS_B}{dt} = (b_B - (b_B - m_B)N_B/K_B)N_B - k(T)p_M\delta_M I_M/K_B S_B - m_B S_B \quad (5)$$

$$\frac{dE_B}{dt} = k(T)p_M\delta_M I_M/K_B S_B - \gamma_B E_B - m_B E_B \quad (6)$$

$$\frac{dI_B}{dt} = \gamma_B E_B - \alpha_B I_B - m_B I_B \quad (7)$$

$$\frac{dR_B}{dt} = (1 - v_B)\alpha_B I_B - m_B R_B \quad (8)$$

$$\frac{dD_B}{dt} = v_B\alpha_B I_B \quad (9)$$

For further details of the model and its calibration see Rubel et al. (2008), where model parameters were set to values based on extensive literature search, then some model tuning was made, and model fit was checked qualitatively. Our aim was to work out an appropriate methodology by which a deterministic model can be transformed into a stochastic one, and to test it by applying

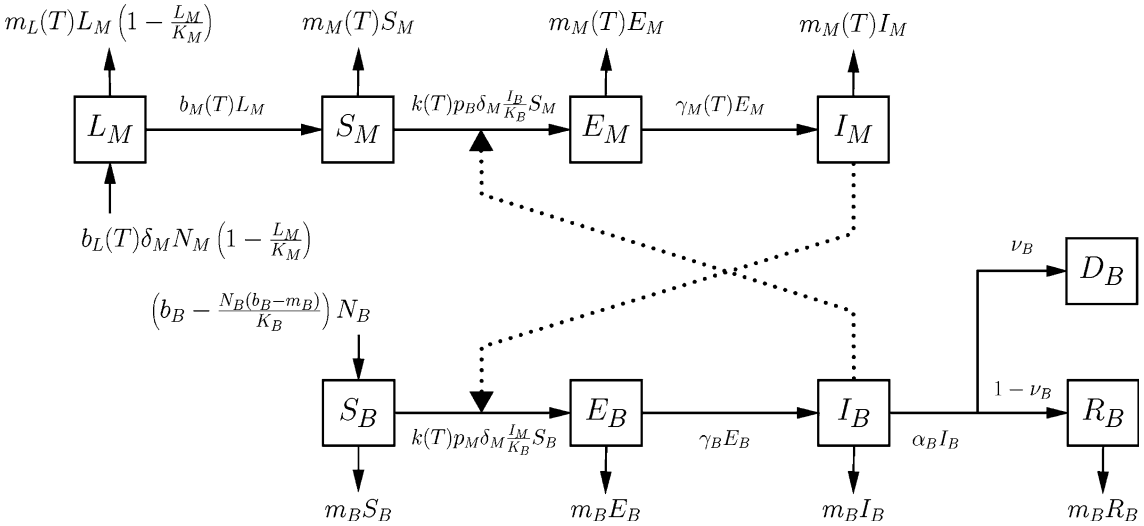

Fig. 2 Diagram of the deterministic SEIR model developed by Rubel et al. (2008). The upper four compartments represent mosquitoes, the lower five compartments represent birds. *Dotted lines* refer to cross-infection

Table 1 Temperature-independent parameters of the model

Parameter	Value
p_B: "Bird to mosquito" transmission probability	0.125
p_M: "Mosquito to bird" transmission probability	1
K_M: Carrying capacity of the environment for mosquitoes (per bird)	100
$N_{M,min}$: Mosquitoes per bird surviving winter	1
m_B: Natural (non-USUV related) mortality rate of birds (per day)	0.0012
α_B: Removal rate (recovery or death) for infected birds (per day)	0.182
γ_B: "Infected to infectious" rate for birds (per day)	0.667
v_B: Probability of death from USUV among infected birds	0.3

it to the above model. We have chosen the hierarchical Bayesian approach (Gelman et al. 2004; Wikle and Berliner 2007), which allows for embedding a deterministic dynamical model in a probabilistic modeling framework (Berliner et al. 2000; Wikle et al. 2001; Wikle 2003). By this methodology we carried out a Bayesian analysis of the eight temperature-independent scalar parameters of the model (Table 1). Analysis of the temperature-dependent vector parameters would require more data.

Several authors pointed out that ignoring the stochastic components in model building may strongly bias the results (Clark et al. 2001; Clark and Gelfand 2006). In our case the most important stochastic features to consider were the following. The observed number of infected dead birds is clearly a random variable. First, not all dead birds were found and sent to the laboratory, and second, the diagnostic test has sensitivity and specificity less than 100%. Daily temperature values also exhibit randomness, varying in space and time. Dependence of the reproduction rates of mosquitoes on temperature is also a stochastic relationship. Same is true for the reproduction rates of birds modeled as a function of calendar day. Other model parameters (birth rates, mortality rates) might also be better described by random variables than by fixed values, since bird and mosquito populations consist of subpopulations with characteristics varying in space and time.

Another advantage of a stochastic model is that in a stochastic model one can quantify the uncertainty in parameter estimates and model predictions in terms of variance estimates or confidence intervals (Clark and Gelfand 2006).

The hierarchical Bayes approach allows for statistical analysis of a model that describes the mechanism of the process through a dynamical system. This can be regarded as an important breakthrough, because in the past there were very few links between dynamical and statistical modeling (West and Harrison 1997; Bolker 2008). Most dynamical models failed at quantification of uncertainty while most statistical models failed at mirroring the real mechanisms of how the system works (Clark and Gelfand 2006).

2 Materials and methods

2.1 Hierarchical Bayesian modeling

The main idea of the hierarchical Bayes approach is to divide a complex process into a hierarchy of submodels or levels, conditional on one another. In such a model stochastic components appearing at different levels are assumed to be independent from each other (Wikle et al. 1998). This way, there is no need to specify a complicated multidimensional prior distribution reflecting all possible interdependencies between the model components and parameters. In most cases this is quite realistic: measurement error in observed data has nothing to do with randomness in the model parameters, nor with randomness in external process variables.

Here we use a hierarchy consisting of three levels: the data model, the process model, and the parameter model. The data model specifies the distribution of data given the process and parameters. The process model describes the process output conditioned on the process parameters. The parameter model accounts for the uncertainty in the process parameters. In other words emphasizing conditionality, process model is conditional on the parameters, while data model is conditional on both the parameters and the process (Berliner 1996; Wikle et al. 1998).

Details of the aforementioned submodels are specified as follows. At the parameter level, we assumed uniform prior distributions for all temperature-independent parameters shown in Table 1. The ranges for these uniform distributions were set initially from half to double of the parameter value applied in the deterministic dynamical model (except probabilities having a bounded range). Later, if MCMC simulations produced paths residing at the boundary of this pre-specified range, we shifted the involved boundaries to allow more space for the paths and repeated the simulation with that extended range. The original values in the deterministic model as well as the final ranges are shown in Table 2. Temperature-dependent features, e.g. reproduction of birds or mosquitoes, were not analyzed because in the dynamical model they were modeled by curves, and so their analysis would require estimation of several parameters per feature. However, the observed data series did not provide enough information to estimate many parameters. Hence, for all temperature-dependent parameters, the same values were used that were applied in the deterministic model.

Table 2 Values of parameters used in the deterministic model and final ranges for the uniform prior distributions in the stochastic model

Parameter	p_B	p_M	K_M	$N_{M,min}$	m_B	α_B	γ_B	v_B
Original value	0.125	1.0	100	1.0	0.0012	0.182	0.667	0.30
Range from	0.050	0.5	2	0.5	0.0001	0.090	0.300	0.01
to	0.300	1.0	200	3.0	0.0060	0.370	1.500	0.70

At the process level, we assumed random fluctuation in temperature data as well as in reproduction numbers of blackbirds. Daily temperature data were perturbed by adding a Gaussian random variable with mean = 0 and SD = 1.25°C. SD was estimated from the data of 12 weather stations provided by the Austrian National Weather Service (Zentralanstalt für Meteorologie und Geodynamik) in or near Vienna. The stations were Baden, Fuchsenbigl, Grossenzersdorf, Gumpoldskirchen, Hohe Warte, Innere Stadt, Langenlebarn, Schwechat, Stockerau, Unterlaa, Wiener Neustadt, and Donaufeld. The reproduction process of blackbirds was made random by multiplying the daily reproduction rates used in the deterministic model by a uniform random number between 0.9 and 1.1, resulting in a random component of about ±10%. Apart from this we used the original dynamical model, implemented in FORTRAN, without any modification.

At the data level, we assumed that the observed number of birds died from USUV is a Poisson variable with mean value (i.e., Poisson parameter) coming from the process model.

The joint posterior distribution of the parameters was computed by Markov Chain Monte Carlo (MCMC) simulation, and posterior means and medians as well as credible intervals were calculated from this distribution. Distribution of the model predictions (that is, predicted number of birds died from USUV) was determined by simulating parameters from their 97.5% posterior intervals independently of each other, resulting in a 90% four-dimensional posterior set, and calculating model predictions for all these simulated parameter combinations.

2.2 Simulation strategy

The analysis by MCMC was not straightforward due to the interdependencies between some model parameters. Rubel et al. (2008) noticed that transmission probabilities p_B and p_M are correlated negatively, but we suspected that this was not the only correlation between the model parameters. Preliminary analysis with one- and two-dimensional log-likelihood maps revealed that the transmission probabilities p_B and p_M, the number of mosquitoes surviving winter $N_{M,min}$, and the removal rate α_B were strongly correlated, moreover, they formed likelihood plateaus (Figs. 3, 4), so putting all of them together into a MCMC simulation could prevent convergence of the chain.

Log-likelihood maps were made as follows. For a single parameter (in this case the map was actually a curve) 50 equidistant points were taken along the prior range of the parameter of interest, log-likelihood was calculated at each point 500 times, and the maximum of these 500 values was taken at each point (for an example see Fig. 3). Two-dimensional log-likelihood maps were obtained similarly for pairs of parameters. Here log-likelihood was calculated 500 times at each point on a 50 × 50 grid, and again, maximum was taken at each point (Fig. 4). We call the resulted log-likelihood map conditional, if all other parameters are kept constant, and unconditional, if all other parameters take random values from their prior distributions.

Due to the revealed likelihood plateaus parameters p_B, p_M, $N_{M,min}$, and α_B cannot be identified in the model, i.e., it is impossible to estimate all of them simultaneously from the observed data. However, if three of them could obtain values from outside (e.g., from laboratory experiments), the fourth parameter could be estimated from the observed data. Therefore, we have chosen those three, namely the

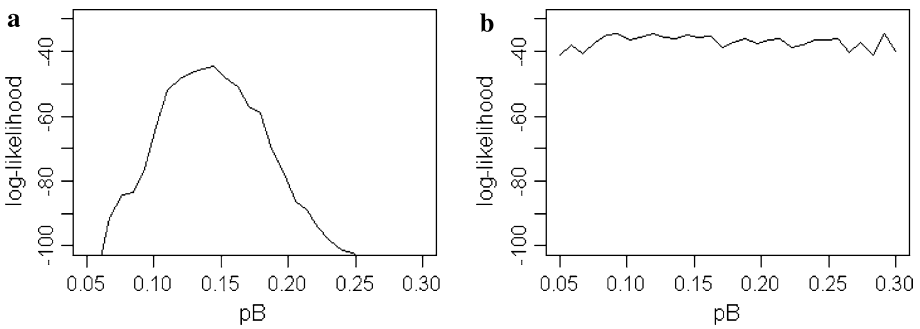

Fig. 3 Conditional log-likelihood curve of the bird to mosquito transmission probability p_B, keeping other parameters at those values used in the deterministic model (**a**); and unconditional log-likelihood curve for the same parameter, randomly varying all other parameters according to their prior distributions (**b**). In each case, function values are maxima of 500 replications, evaluated at 50 equidistant points along the prior range of p_B

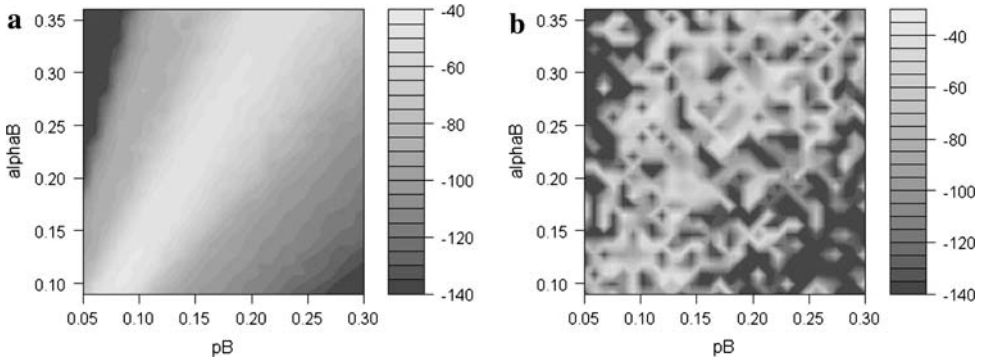

Fig. 4 Conditional (**a**) and unconditional (**b**) log-likelihood map of the bird to mosquito transmission probability p_B versus removal rate α_B. In each case, function values are maxima of 500 replications at each point on a grid of 50 × 50 over the joint prior range of p_B and α_B

transmission probabilities p_B, p_M, and the removal rate α_B, which can be—at least theoretically—determined in laboratory experiments similar to those by Weissenböck et al. (2004), Bakonyi et al. (2005), Chvala et al. (2005). To these parameters we assigned the same values that were used in the deterministic model, namely we took $p_B = 0.125$, $p_M = 1$, $\alpha_B = 0.182$, while the number of mosquitoes surviving winter, $N_{M,min}$ was determined by the MCMC simulation described above. To find out whether there are combinations of p_B, p_M, and α_B resulting in higher likelihood than the one used in the deterministic model, we made log-likelihood maps of $N_{M,min}$ for 27 further combinations of the first three parameters (not reported). We found that the log-likelihood obtained for the above parameter combination was not substantially exceeded by other combinations (differences were comparable with random variability).

Another difficulty was experienced when we tried to include the "infected to infectious" rate, γ_B into the model because this prevented convergence of the MCMC. Therefore, we made log-likelihood maps with γ_B, and we found that the likelihood did not depend on γ_B (Fig. 5). The reason could be that observed data did not provide any information on γ_B, since no data related to compartments E_B and I_B were available. Hence in the final model γ_B was fixed at the value used by Rubel et al. (2008).

2.3 Markov chain Monte Carlo simulation

The MCMC method serves for estimation of posterior densities of parameters in Bayesian models (Gelman and Rubin 1996; Congdon 2006). Its application area has been continuously expanding since its first appearance (Aldstadt 2007; Muller et al. 2008). In the present study MCMC simulation was achieved by a random walk Metropolis algorithm with 300,000 iterations. Running time was about 10 h on a PC with an 1.8 GHz AMD Sempron 3100 processor, 1.5 GB RAM, Windows 2000, R 2.4.1, Compaq Visual Fortran v6.1. The MCMC algorithm was programmed by the first author in R, calling the FORTRAN subroutine for the original dynamical model from R via the .Fortran() function. This ensured that our process model was truly identical with the one used by Rubel et al. (2008).

Jumping distribution of the Metropolis algorithm was specified to be uniform with a range of ±5% of the range of the prior distributions in Table 2. (i.e., ±0.125 for $N_{M,min}$, ±0.0003 for m_B, etc.). In each step all parameters were updated simultaneously according to the jumping distribution. When a jump resulted in a point outside the parameter range, the point was replaced by another point within the range at the same distance but from the other end. For example, if the parameter range was $[a, b]$ and a jump

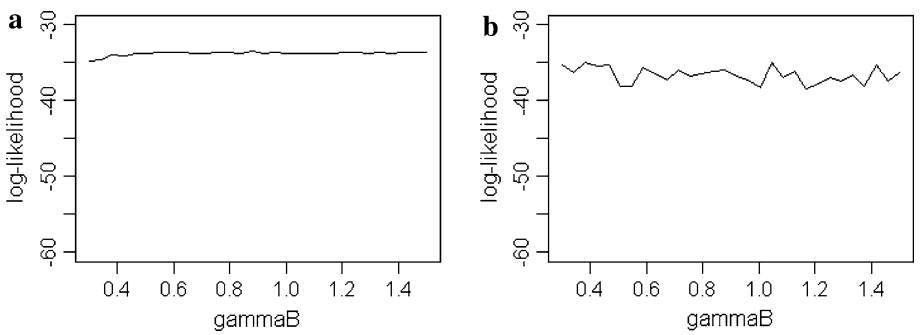

Fig. 5 Conditional (**a**) and unconditional (**b**) log-likelihood curves of the "infected to infectious" rate γ_B, showing that γ_B has no influence on the likelihood. In each case, values are maxima of 500 replications, evaluated at 50 equidistant points along the prior range of γ_B

Table 3 Initial parameter values for the three chains

	$N_{M,min}$	K_M	m_B	v_B
Chain 1:	1.0	50	0.0021	0.4
Chain 2:	1.8	20	0.0040	0.1
Chain 3:	0.6	150	0.0008	0.6

resulted in $b + u$, then it was replaced by $a + u$. This method preserved the symmetry of the proposal distribution.

Convergence of the Markov chain was checked using the Gelman–Rubin scale reduction factor (Gelman and Rubin 1992, 1996; Congdon 2006) and the Brooks–Gelman–Rubin statistic (Brooks and Gelman 1998; Congdon 2006), each one calculated from three independent chains. Initial parameter values for the three chains (Table 3) were selected according to the recommendation of Congdon (2006). Autocorrelation was checked graphically.

To see how uncertainty in the parameters affects uncertainty of predictions, we made a simulation with the final model. We drew 10,000 random parameter combinations from their 97.5% posterior intervals, resulting in an approximate 90% four-dimensional joint confidence set.

3 Results

The final model included four parameters to analyze while the others were kept constant at the values proposed by Rubel et al. (2008). The studied parameters were the number of mosquitoes surviving winter $N_{M,min}$, the carrying capacity of the environment for mosquitoes K_M, the natural (non-USUV-related) mortality rate of birds m_B, and the proportion of infected birds dying from USUV v_B. Posterior means, medians, 90% and 95% posterior intervals for these parameters are summarized in Table 4. Estimates are calculated from three independent MCMC runs with different initial states (see Table 3), each with 300,000 iterations. From each run the last 150,000 iterations were used (assuming a burn-in period of 150,000 iterations), so each estimate in Table 4 is based on $3 \times 150{,}000 = 450{,}000$ values.

The conditional and unconditional log-likelihood curves of the "infected to infectious" rate γ_B are shown in Fig. 5. The conditional curve is based on the values used in the deterministic model for p_B, p_M, and α_B together with the posterior median values for $N_{M,min}$, K_M, m_B, and v_B. The unconditional curve is based on 500 random parameter combinations from the prior distributions at each of the 50 points.

The acceptance rates for the three runs were 33.8%, 33.9%, and 26.4%. Convergence results are summarized in Table 5. Both the Gelman–Rubin potential scale reduction factor (PSRF) and the Brooks–Gelman–Rubin statistic (BGR) were calculated assuming a burn-in period of 150,000 iterations. Results supported convergence of the chains.

Autocorrelation in the Markov chains was quite high, thinning factors 50, 50, 100, and 300 were needed to eliminate autocorrelation in $N_{M,min}$, K_M, m_B, and v_B, respectively.

Figure 6 shows the variability of the number of birds dead from USUV, as predicted by the model if varying the parameters in the 97.5% posterior intervals shown in Table 4. Figure 6 is based on a simulation with 10,000 random parameter combinations from the posterior intervals. The diagram shows that, although the model fits acceptably to observed data, the predicted dead bird counts cover a rather wide range for 2002 and 2003.

4 Discussion

Spread of infectious diseases to new geographical areas during the last decades induced several attempts using various approaches to model outbreak dynamics of infectious diseases (Mikler et al. 2007; Aldstadt 2007; Choi et al. 2008). The present study focused on the question how to transform a deterministic dynamic model into a stochastic one. Embedding the deterministic model into a hierarchical Bayes model enabled taking into account some stochastic features of the process and to quantify the uncertainty in parameters and predictions. At the same time, Bayesian analysis gave deeper insight in the structure of the model. We experienced that certain parameters were not identifiable due to redundancy in the parameterization of the model, and one further parameter did not influence the likelihood. Fortunately, none of these findings discredited the model since all of them had clear interpretations with respect to the epidemiological process.

Table 4 Posterior means, medians, and credible intervals for the parameters

Parameter	$N_{M,min}$	K_M	m_B	v_B
Posterior mean	1.37	18.0	0.0029	0.233
Posterior median	1.37	16.1	0.0027	0.191
90% posterior (credible) interval	(1.18, 1.58)	(9.13, 33.2)	(0.0010, 0.0052)	(0.038, 0.573)
95% posterior (credible) interval	(1.15, 1.62)	(8.38, 39.1)	(0.0007, 0.0056)	(0.025, 0.630)

Table 5 Gelman–Rubin potential scale reduction factor (PSRF) and Brooks–Gelman–Rubin statistic (BGR) calculated from three independent chains, assuming a burn-in period of 150,000 iterations

	$N_{M,min}$	K_M	m_B	v_B
PSFR	1.0001	1.0002	1.0008	1.0008
BGR	1.0016	0.9948	1.0050	1.0026

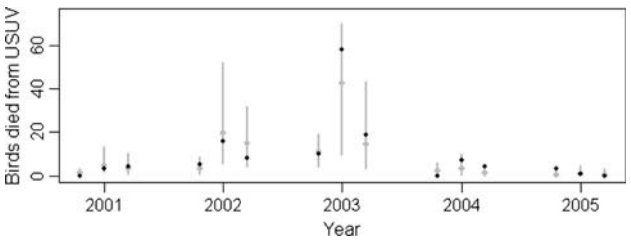

Fig. 6 Predicted (*grey*) and observed (*black*) number of birds died from USUV. *Grey diamonds* denote the mean value of predicted numbers. The diagram is based on 10,000 predicted data series, varying the model parameters in their 97.5% posterior intervals independently of each other

We found that γ_B, i.e., the rate "infected to infectious" (related to the length of the intrinsic incubation period) did not influence the likelihood at all, which means that this parameter was not identifiable in the model. This is not surprising because observed data are restricted to one single compartment, namely to that of dead birds, furthermore they are monthly aggregates. As the intrinsic incubation period is rather short (a few days), it causes just a short delay in the transmission of the disease. During such a short period neither weather conditions nor population sizes (neither for birds nor for mosquitoes) can change considerably, so this short delay cannot have a substantial effect on the number of dead birds.

It is also quite reasonable that the transmission probabilities p_B and p_M are negatively correlated because both of them influence the speed of spread positively. The fact that the likelihood is about the same in a large region in the p_B, p_M plane means that p_B and p_M cannot be estimated independently of each other. One can give a conditional posterior interval for one of them conditioning on the other one, or one can give a joint (two-dimensional) posterior range. This also means that if one of them could be determined in a laboratory experiment, then the other one could be estimated from the model.

Further strong—nearly linear—relationships were found between α_B, i.e., the removal rate of infected birds (towards recovery or death), and the transmission probabilities p_B and p_M. The reason for this may be that the removal rate is inversely proportional to the time span of the disease, i.e., to the time from infection to cure or death. In the light of this, the experienced relationship is quite reasonable intuitively, because assuming a longer disease period but smaller transmission probabilities may result in nearly the same epidemiological curve.

The carrying capacity of environment for mosquitoes K_M is a parameter reflecting the population size and/or the reproduction patterns of mosquitoes, and presumably depends on meteorological conditions. Its estimated value deviates most considerably from its value in the initial deterministic SEIR model. This suggests that further investigation is needed concerning the reproduction of mosquitoes, or some estimates are needed on what proportion of the mosquito population is involved in USUV transmission.

Parameter m_B, the non-USUV mortality rate of birds, shows the second largest deviation from its value in the initial model. Although the value is contained in the 90% posterior interval, it is also possible that the literature has not reflected to the latest changes in the life span of blackbirds yet. Estimates of other parameters are fairly close to the values used in the deterministic model (see Tables 2, 4): posterior mean of $N_{M,min}$ has become greater while that of v_B smaller but the differences do not seem to be of much relevance. Further versions of the model should also reflect to the fact that $N_{M,min}$ is a temperature-dependent parameter that is likely to depend on the winter temperature.

One could speculate that if longer data series were available, more parameters could be estimated, but due to the interplay between the parameters it is more likely that longer data series could not help either. In accordance with this, we decided to carry out a conditional inference by fixing the transmission probabilities p_B, p_M, and the removal rate α_B, and analyzing the other parameters conditionally on these. To reduce the redundancy in the parameterization of the model, it would be possible to combine the above parameters into a single identifiable "bird to bird" transmission parameter, but we had a good reason for not doing so: the model would loose its power of describing the mechanism of spread of disease. In particular, we would lose control on the population dynamics of mosquitoes, which is an important part of the model, being most sensitive to weather conditions.

An important difference between deterministic and stochastic modeling lies in the assessment of how well the model fits to data. In a deterministic model there is only one single curve and one single observed data series, so the only possibility to quantify the goodness-of-fit is to use a "distance measure" like the mean square error (*MSE*). In the Bayesian framework the likelihood is used instead. These two measures are different both numerically and regarding their interpretation. It may well occur that different parameters have equal likelihood but result in quite different *MSE*, or vice versa. Note that *MSE* relies on the absolute differences between predicted and observed

values, while in our case the likelihood approach, as one based on the Poisson distribution, allows for larger differences if the values themselves are large. To illustrate that Bayesian parameter estimates perform well even in the deterministic context, we calculated *MSE* putting the posterior medians of the parameters into the deterministic model. This resulted in $MSE = 9.91$ compared to $MSE = 16.91$ of the original deterministic model, which is a considerable improvement even in terms of *MSE*.

As a conclusion, the hierarchical Bayes approach proved to be fruitful in extending a deterministic model into a stochastic one. This methodology allowed for statistical analysis, such as quantification of uncertainty in parameters and predictions, as well as Bayesian point and interval estimation. The main difficulties were related to identifiability of some parameters, but this is in a certain sense unavoidable: one must always find a compromise between models that describe the mechanism of the process as precise as possible, but probably exhibit interrelationships between their parameters (which is most likely so in the real system too), and models with identifiable parameters that have less natural interpretation with respect to the modeled real system.

Acknowledgments This research was made during the first author's stay at the University for Veterinary Medicine, Vienna in the frame of a Marie Curie Inter-European Fellowship (Applied stochastic modeling in veterinary epidemiology, contract no. 39193). The work was also supported by the Hungarian National Research Fund, grant no. OTKA T049157. Temperature data were provided by the Austrian Weather Service (ZAMG).

References

Aldstadt J (2007) An incremental Knox test for the determination of the serial interval between successive cases of an infectious disease. Stoch Environ Res Risk Assess 21:487–500

Bakonyi T, Lussy H, Weissenböck H, Hornyák A, Nowotny N (2005) In vitro host-cell susceptibility to Usutu virus. Emerg Infect Dis 11:298–301

Bakonyi T, Erdélyi K, Ursu K, Ferenczi E, Csörgo T, Lussy H, Chvala S, Bukovsky C, Meister T, Weissenböck H, Nowotny N (2007) Emergence of Usutu virus in Hungary. J Clin Microbiol 45:3870–3874

Berliner LM (1996) Hierarchical Bayesian time-series models. In: Hanson K, Silver R (eds) Maximum entropy and Bayesian methods 15–22. Kluwer, Dordrecht

Berliner LM, Wikle CK, Cressie N (2000) Long-lead prediction of Pacific SSTs via Bayesian dynamic modelling. J Clim 13(22): 3953–3968

Bolker BM (2008) Ecological models and data in R. Princeton University Press, Princeton, p 396

Brooks S, Gelman A (1998) General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 7: 434–456

Brugger K, Rubel F (2009) Simulation of climate-change scenarios to explain Usutu-virus dynamics in Austria. Prev Vet Med 88:24–31

Choi KM, Yu HL, Wilson ML (2008) Spatiotemporal statistical analysis of influenza mortality risk in the State of California during the period 1997–2001. Stoch Environ Res Risk Assess 22(Suppl.1):S15–S25

Chvala S, Bakonyi T, Hackl R, Hess M, Nowotny N, Weissenböck H (2005) Limited pathogenicity of Usutu virus for the domestic chicken (*Gallus domesticus*). Avian Pathol 34:392–395

Chvala S, Bakonyi T, Bukovsky C, Meister T, Brugger K, Rubel F, Nowotny N, Weissenböck H (2007) Monitoring of Usutu virus activity and spread by using dead bird surveillance in Austria, 2003–2005. Vet Microbiol 122:237–245

Clark JS, Gelfand AE (2006) A future for models and data in environmental science. Trends Ecol Evol 21(7):375–380

Clark JS, Carpenter SR, Barber M, Collins S, Dobson A, Foley JA, Lodge DM, Pascual M, Pielke R Jr, Pizer W, Pringle C, Reid WV, Rose KA, Sala O, Schlesinger WH, Wall DH, Wear D (2001) Ecological forecasts: an emerging imperative. Science 293: 657–660

Congdon P (2006) Bayesian statistical modelling, 2nd edn. Wiley, New York

Gelman A, Rubin D (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–511

Gelman A, Rubin D (1996) Markov chain Monte Carlo methods in biostatistics. Stat Methods Med Res 5:339–355

Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis, 2nd edn. Chapman and Hall, CRC, Boca Raton

Keeling MJ, Rohani P (2007) Modeling infectious diseases in humans and animals. Princeton University Press, Princeton, p 408

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World map of Köppen-Geiger climate classification updated. Meteorol Z 15: 259–263

Mikler AR, Venkatachalam S, Ramisetty-Mikler S (2007) Decisions under uncertainty: a computational framework for quantification of policies addressing infectious disease epidemics. Stoch Environ Res Risk Assess 21:533–543

Muller A, Bacro JN, Lang M (2008) Bayesian comparison of different rainfall depth–duration–frequency relationships. Stoch Environ Res Risk Assess 22:33–46

Rubel F, Brugger K, Hantel M, Chvala-Mannsberger S, Bakonyi T, Weissenböck H, Nowotny N (2008) Explaining USUTU virus dynamics in Austria: model development and calibration. Prev Vet Med 85:166–186

Weissenböck H, Kolodziejek J, Url A, Lussy H, Rebel-Bauder B, Nowotny N (2002) Emergence of Usutu virus, an African mosquito-borne flavivirus of the Japanese encephalitis virus group, central Europe. Emerg Infect Dis 8:652–656

Weissenböck H, Bakonyi T, Chvala S, Nowotny N (2004) Experimental Usutu virus infection of suckling mice causes neuronal and glial cell apoptosis and demyelination. Acta Neuropathol 108:453–460

West M, Harrison J (1997) Bayesian forecasting and dynamic models, 2nd edn. Springer, New York, p 680

Wikle CK (2003) Hierarchical models in environmental science. Int Stat Rev 71:181–199

Wikle CK, Berliner LM (2007) A Bayesian tutorial for data assimilation. Physica D 230(1–2):1–16

Wikle CK, Berliner LM, Cressie N (1998) Hierarchical Bayesian space-time models. Environ Ecol Stat 5(2):117–154

Wikle CK, Milliff RF, Nychka D, Berliner LM (2001) Spatiotemporal hierarchical Bayesian modeling: tropical ocean surface winds. J Am Stat Assoc 96:382–397