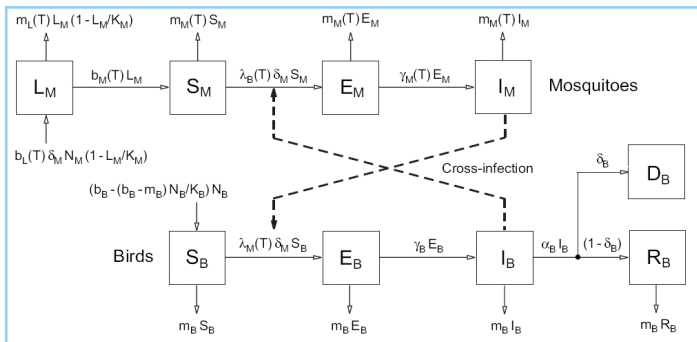


Introduction

The USUTU virus is an arbovirus transmitted by mosquitos and causing disease in blackbirds. The virus was first detected in Austria in 2001, and a major outbreak occurred in 2003. The histogram shows the numbers of USUV-positive dead birds found in Eastern Austria from 2001 to 2005. Monitoring took place each year only in July, August, and September.

Rubel et al. [1] developed a 9-compartment SEIR model to explain the spread of disease (see the diagram below). The model has several parameters, some of which are temperature-dependent. Rubel et al. [1] set the model parameters to values taken from the literature, made some model tuning, and checked the model fit qualitatively.

Our aim was to work out a methodology by which a deterministic model can be transformed into a stochastic one, and to apply this to the above model. To do this, we used the hierarchical Bayes approach [2], and carried out a Bayesian analysis of the eight temperature-independent scalar parameters of our model. (Analysis of the temperature-dependent vector parameters would require more data.)



Hierarchical Bayes Models

The basic idea: divide the process into parts to obtain a hierarchy of submodels conditional on one another [2]. The simplest structure consists of three levels of hierarchy:

- The *data model* specifies the distribution of the data given the process and its parameters.
- The *process model* describes the process output conditioned on model parameters and on further exogenous variables.
- The *parameter model* accounts for the uncertainty in the process parameters.

The main advantage: there is no need to assume a rather complicated multidimensional prior distribution reflecting to all possible interdependencies between the model components and parameters. Priors at different levels are assumed to be independent.

Statistical analysis can be made by Markov Chain Monte Carlo (MCMC).

Application to the USUTU model

At the data level, we assumed that the observed number of USUV-positive dead birds is a Poisson-distributed random variable with mean value coming from the process model.

At the process level, we used the original dynamical model without any modification, just assumed randomness in some environmental variables, such as in temperature data and 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 near Vienna. The reproduction process of blackbirds was made random by multiplying the daily reproduction rate used in the deterministic model by a uniform random number between 0.9 and 1.1, resulting in a random component of about ±10%.



At the parameter level, we assumed uniform priors for all temperature-independent model parameters, namely for p_B , p_M , K_L , $N_{M,min}$, m_B , α_B , γ_B , and δ_B (for explanation see the summary table below). The ranges for the uniform priors were set initially from the half to the double of the parameter value applied in the deterministic dynamical model (except for probabilities having a bounded range). Later if MCMC simulations produced paths residing at the boundary of this pre-specified range, we shifted the corresponding boundary to give more place to the paths.

Details of the MCMC simulation

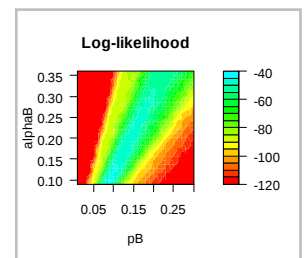
MCMC was done by a random walk Metropolis-Hastings algorithm with 200 000 iterations. Jumping distribution was uniform with a range of ±5% of the given prior range for each parameter. Jumps outside the prior range were not accepted.

The model was implemented in R, combined with a Fortran subroutine calculating the original deterministic dynamical model.

Convergence of the Markov chain was checked using the Gelman-Rubin scale reduction factor and the Brooks-Gelman-Rubin statistic [3] calculated from 3 independent chains. Autocorrelation was checked graphically.

Results

The analysis by MCMC was not straightforward due to the interdependencies between some model parameters. Analysis of likelihood maps revealed that parameters p_B , p_M , $N_{M,min}$, and α_B are strongly correlated, forming likelihood plateaus, so putting all of them together into a MCMC simulation would prevent convergence of the chain. The diagram illustrates the dependency between p_B and α_B .



Therefore we fixed the parameters p_B , p_M , and α_B at the values used in the deterministic SEIR model, and determined the posterior distribution of the others conditional on these. Now the MCMC algorithm converged except for γ_B , which does not seem to have any influence on the likelihood. Initial parameter values for the 3 chains used for convergence assessment are as follows:

	$N_{M,min}$	K_L	m_B	δ_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

Results from the MCMC simulation are shown in the summary table below.

References:

- [1] 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. Accepted for publication in *Preventive Veterinary Medicine*.
- [2] Wikle CK (2003) Hierarchical models in environmental science. *International Statistical Review*, 71, 181-199.
- [3] Congdon P (2006) *Bayesian Statistical Modelling* (2nd ed.), Wiley

This research was partly supported by the grant OTKA T049157.

Summary information about the parameters (columns 1 to 3), and results from the MCMC simulation (columns 4 to 6)

Parameter	Value in orig. model	Prior range	Post. mean	Post. median	95% post. interval
p_B : „Bird to mosquito” transmission probability	0.125	(0.05, 0.30)	♦	♦	♦
p_M : „Mosquito to bird” transmission probability	1	(0.5, 1)	♦	♦	♦
K_L : Carrying capacity of environment for mosquito larvae	100	(2, 200)	18.6	16.4	(8.47, 41.5)
$N_{M,min}$: Mosquitos per bird surviving winter	1	(0.5, 3)	1.37	1.36	(1.14, 1.62)
m_B : Natural (non-USUV related) mortality rate of birds (per day)	0.0012	(0.0001, 0.0060)	0.0029	0.0028	(0.0007, 0.0055)
α_B : Removal rate (recovery or death) for infected birds (per day)	0.182	(0.09, 0.37)	♦	♦	♦
γ_B : „Infected to infectious” rate for birds (per day)	0.667	(0.3, 1.5)	♣	♣	♣
δ_B : Fraction of infected birds dying from USUV	0.3	(0.01, 0.7)	0.222	0.182	(0.023, 0.610)

♦ not estimable due to interdependencies between parameters; ♣ not estimable (likelihood function is constant)