

Cholesky Decomposition for the Vasicek Interest Rate Model

Muhannad Al-Saadony¹, Paul Hewson² & Julian Stander²

¹ Department of Statistics, Al-Qadisiyah University, Al-Qadisiyah, Iraq

² School of Computing & Mathematics, Plymouth University, UK

Correspondence: Muhannad Al-Saadony, Department of Statistics, Al-Qadisiyah University, Al-Qadisiyah, Iraq.
E-mail: muhannad.alsaadony@gmail.com

Received: June 23, 2013 Accepted: July 14, 2013 Online Published: August 20, 2013

doi:10.5539/ijsp.v2n4p22 URL: <http://dx.doi.org/10.5539/ijsp.v2n4p22>

Abstract

This paper concerns the estimation of parameters in the “Vasicek Interest Rate” model under a Bayesian framework. These popular models are challenging to fit with Markov chain Monte Carlo (MCMC) methods as the structure of the model leads to considerable autocorrelation in the chains. Accordingly, we demonstrate that a simple reparameterisation using the Cholesky decomposition can greatly improve the performance of the MCMC algorithm and hence lead to valid Bayesian inference on the Vasicek model.

Keywords: Vasicek interest rate model, maximum likelihood estimation, MCMC method, autocorrelation, Cholesky decomposition

1. Introduction

The “Vasicek Interest Rate” model (referred to subsequently as the Vasicek model) is a popular example of a stochastic differential equation that models interest rate as a function of market risk (Vasicek, 1977). The Vasicek model remains popular and there is considerable current interest in this model; for example Chua et al. (2013) report on approaches using Bayesian model averaging comparing high frequency and weekly data. This is despite a potential flaw of the model as originally proposed in that it can allow for a negative interest rate. Consequently later developments on this model include for example the Cox-Ingersoll-Ross (CIR) model and the Hull-White model (Jacus, 2008) but as we note later these models share common features with the Vasicek model and could benefit from the proposals demonstrated in this paper. As usual, the challenge with the Bayesian approach is parameter estimation. We therefore use Markov chain Monte Carlo (MCMC) methods to simulate from posterior distribution of the parameters in our model (Cowles & Carlin, 1996; Gilks et al., 1996). This in turn creates a problem in that the posterior simulations have high autocorrelation due largely to the specification of the model. Hence we propose to use a Cholesky decomposition (Gene & Charles, 2013) to reparameterise the model in a way that reduces this structural autocorrelation and consequently enhances the ease with which we can perform Bayesian inference on this model.

The paper is organized as follows; in Section 2 we present the Vasicek model outlining the likelihood for the model and indicating how we use MCMC to simulate from the posterior distribution of the parameters of the model. In Section 2.2, we show how the Cholesky decomposition can be applied to the parameters of the Vasicek model and how we simulate from the posterior distribution of these parameters. In Section 3 we present results from simulation studies showing that the reparameterisation does indeed reduce autocorrelation and hence we conclude our paper in Section 4.

2. Vasicek Interest Rate Model

The Vasicek Interest Rate model was introduced by Vasicek (1977). It models the instantaneous interest rate r_t by a stochastic differential equation so that

$$dr_t = \theta_1(\theta_2 - r_t)dt + \theta_3 dW_t \quad \theta_1, \theta_2, \theta_3 > 0, \quad (1)$$

where:

r_t is an interest rate process at continuous time t ,

θ_1 is a parameter describing the speed of mean reversion,

θ_2 is a parameter describing the long run mean interest rate,

θ_3 is a parameter describing the instantaneous short-rate volatility and

dW_t is a standard Brownian motion increment over time interval dt .

We note that both the drift and volatility are assumed to be stochastic processes. Hence we wish to determine the transition density; this relies on previous events, i.e. $\varphi(r_t|r_{t-1})$. Denote the likelihood function $L(\theta|r)$ in a discrete time, with parameter $\theta = [\theta_1, \theta_2, \theta_3]$ and data $r = (r_{1\delta}, r_{2\delta}, \dots, r_{n\delta})$, which in full takes the form:

$$\begin{aligned} L(\theta|r) &= \prod_{j=1}^n \varphi(r_{j\delta}|r_{(j-1)\delta}) \\ &= \left[\frac{\pi\theta_3}{\theta_1} \{1 - \exp(-2\theta_1\delta)\} \right]^{-\frac{n}{2}} \times \exp \left(\frac{-\theta_1 \sum_{j=1}^n [r_{j\delta} - r_{(j-1)\delta} \exp(-\theta_1\delta) - \theta_2 \{1 - \exp(-2\theta_1\delta)\}]^2}{\theta_3 \{1 - \exp(-2\delta)\}} \right), \end{aligned} \quad (2)$$

where δ has been incorporated in order to represent the Vasicek model in a discrete time; in our case we take $\delta = 1$.

One potential disadvantage with the Vasicek model is that the interest rate will be negative if the interest rate r_t is greater than the long run mean interest rate parameter θ_2 , i.e., if $r_t > \theta_2$. As noted in Section 1 this has been addressed by proposing extensions to the Vasicek, for example the CIR model mentioned above has $dr_t = \theta_1(\theta_2 - r_t)dt + \theta_3 \sqrt{r_t}dW_t$ so that if $r_0 > 0$ then $\theta_1\theta_2 > \frac{1}{2}\theta_3^2$. In principle this is a simple extension to the basic Vasicek model requiring no additional parameters and we suggest that the material presented in this paper would also be relevant to this model.

2.1 Model Fitting with Metropolis-Hastings Algorithm

For now, we note that in order to define the posterior distribution of the three parameters $p(\theta|r)$ we need both the likelihood given above as well as a prior distribution $p(\theta)$ so that

$$p(\theta|r) \propto L(\theta|r)p(\theta) \quad (3)$$

This posterior density is analytically intractable but it is simple to define a simulation from this distribution using MCMC methodology. We use the Metropolis Hastings algorithm (Robert & Casella, 2009) which requires that we:

- 1) Simulate a candidate value θ_j^* from a proposal density $K(\theta_j^*|\theta_j^{(t-1)})$ for $j = 1, 2, 3$.
- 2) Compute the ratio.

$$R = \frac{p(\theta_j^*|r, \theta_{-j})K(\theta_j^{(t-1)}|\theta_j^*)}{p(\theta_j^{(t-1)}|r, \theta_{-j})K(\theta_j^*|\theta_j^{(t-1)})}$$

(We use the θ_j notation to indicate that we compute the acceptance probability conditional for θ_j conditional on the other currently accepted values of θ . This is done cyclically, for example we calculate the posterior for θ_1^* conditional on θ_2^{t-1} and θ_3^{t-1} , we calculate the posterior for θ_2^* conditional on θ_1^* and θ_3^{t-1} and then θ_3^* conditional on θ_1^* and θ_2^*).

- 3) Compute the acceptance probability $\alpha = \min[R, 1]$.

- 4) Sample a value $\theta_j^{(t)}$ such that $\theta_j^{(t)} = \theta_j^*$ with probability α ; otherwise set $\theta_j^{(t)} = \theta_j^{t-1}$.

The challenge here is that because of the nature of our model, we necessarily induce a high autocorrelation between successive simulated values of θ_1, θ_2 and θ_3 . One solution is to use a Cholesky decomposition to reparameterise the model so that we have three parameters that are no longer themselves correlated thus reducing the autocorrelation in the posterior simulations.

2.2 The Cholesky Decomposition

Our goal is to reduce the correlation between successive simulations from the posterior distribution, which is challenging to achieve given that the posterior distribution of the parameters is itself correlated. Denote an estimate of the posterior variance-covariance matrix of θ .

$$\widehat{\text{var}}[\theta|r] = V$$

where $\theta = [\theta_1, \theta_2, \theta_3]$. For some suitable matrix M , we have that

$$\widehat{\text{var}}[M[\theta]|r] = MVM^T.$$

Hence, if we select M such that $MVM^T = I$ and then fit a model to $M[\theta]$ we have defined a posterior distribution for the parameters which is uncorrelated. One specific solution to this problem is to use the Cholesky decomposition such that $U^T U = V$ and hence we use $M = (U^T)^{-1}$ so that:

$$\begin{aligned} MVM^T &= MU^T U M^T \\ &= (U^T)^{-1} U^T U ((U^T)^{-1})^T \\ &= (U^T)^{-1} U^T U U^{-1} \\ &= I \end{aligned} \quad (4)$$

Therefore, we apply this method to the Vasicek model as follows:

- 1) Estimate the parameters θ and their variance-covariance matrix V .
- 2) Compute transformed parameters $\beta = [\beta_1, \beta_2, \beta_3]$ using

$$\beta = V^{-1}[\theta] \quad (5)$$

where $\beta = [\beta_1, \beta_2, \beta_3]$ are new parameters formed by rotating the original parameters $\theta = [\theta_1, \theta_2, \theta_3]$.

To apply to our model, we can proceed as follows:

- 1) Estimate θ^0 and V^0 using maximum likelihood.
- 2) Find the new parameters β using Equation (5).
- 3) Simulate independent candidate value β_j^* from a proposal density $K(\beta_j^*|\beta_j^{(t-1)})$ for $j = 1, 2, 3$.
- 4) Compute the new proposal for model parameters $\theta^* = U\beta$ given a proposal for one of the β_j 's but leaving the other two β_j 's unchanged.
- 5) Compute the ratio.

$$R = \frac{p(\theta^*|r)K(\beta_j^{(t-1)}|\beta_j^*)}{p(\theta^{(t-1)}|r)K(\beta_j^*|\beta_j^{(t-1)})}$$

- 6) Accept the parameters β_j with probability $\alpha = \min[R, 1]$.

3. Simulation Study

For the standard Vasicek algorithm we select candidate values of θ_1, θ_2 from a log-normal distribution, and candidate values of θ_3 from an inverse-gamma distribution. Likewise, for the transformed parameter algorithm we select candidate values for β_1 and β_2 from a log-Normal distribution and β_3 also from an inverse-gamma distribution. The initial values of the parameters are $\theta = [3, 1, 2]$ and the initial value of the interest rate is 10, in the original model. The initial values of the parameters are $\theta = [3.834, 0.7533, 1.3695]$ and the initial value of the interest rate is 6.22, in the modified model. The prior distributions of θ_1 and β_1 are a gamma distribution, for θ_2 and β_2 are a normal distribution and for θ_3 and β_3 are an inverse gamma. We noticed that our chosen prior is the best prior distributions for all parameter because these distributions give us the best results. We will choose some values of $\sigma = [\sigma_1, \sigma_2, \sigma_3]$ in the interval $[0, 1]$. We noticed that our values of σ is the best choice because a small values of σ is to converge to the correct parameters values as shown in (Everitt, 2009).

We therefore simulate realizations θ and β from their posterior distribution using 10,000 iterations of the Metropolis-Hastings algorithm. Our results are presented graphically using time series plots of the parameter, autocorrelation functions, and scatter-plots of simulated parameter values in Figure 1 for the original model and Figure 2 for the modified model. There is a clear indication of the reduction in autocorrelation using the reparameterised models from these visual results.

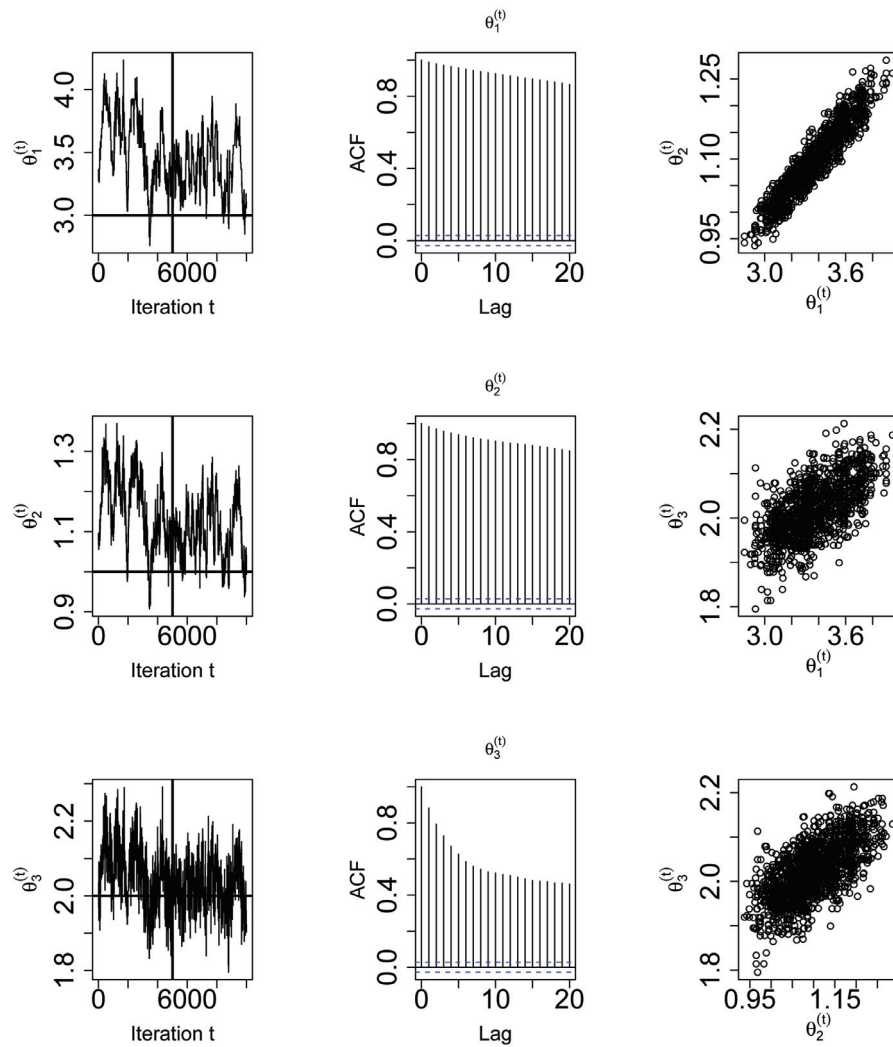


Figure 1. Results from a MCMC algorithm for sampling from the posterior $p(\theta|r)$ where $\theta = [\theta_1, \theta_2, \theta_3]$. The first column shows time series plots of sampled values $\theta_1^{(t)}$, $\theta_2^{(t)}$ and $\theta_3^{(t)}$. The vertical lines separates the burn-in phase(left) from the samples that used for future inference(right). Partial autocorrelation function are shown in the second column. Finally, $\theta_1^{(t)}$ against $\theta_2^{(t)}$, $\theta_2^{(t)}$ against $\theta_1^{(t)}$, $\theta_3^{(t)}$ against $\theta_2^{(t)}$ are shown in the the third column

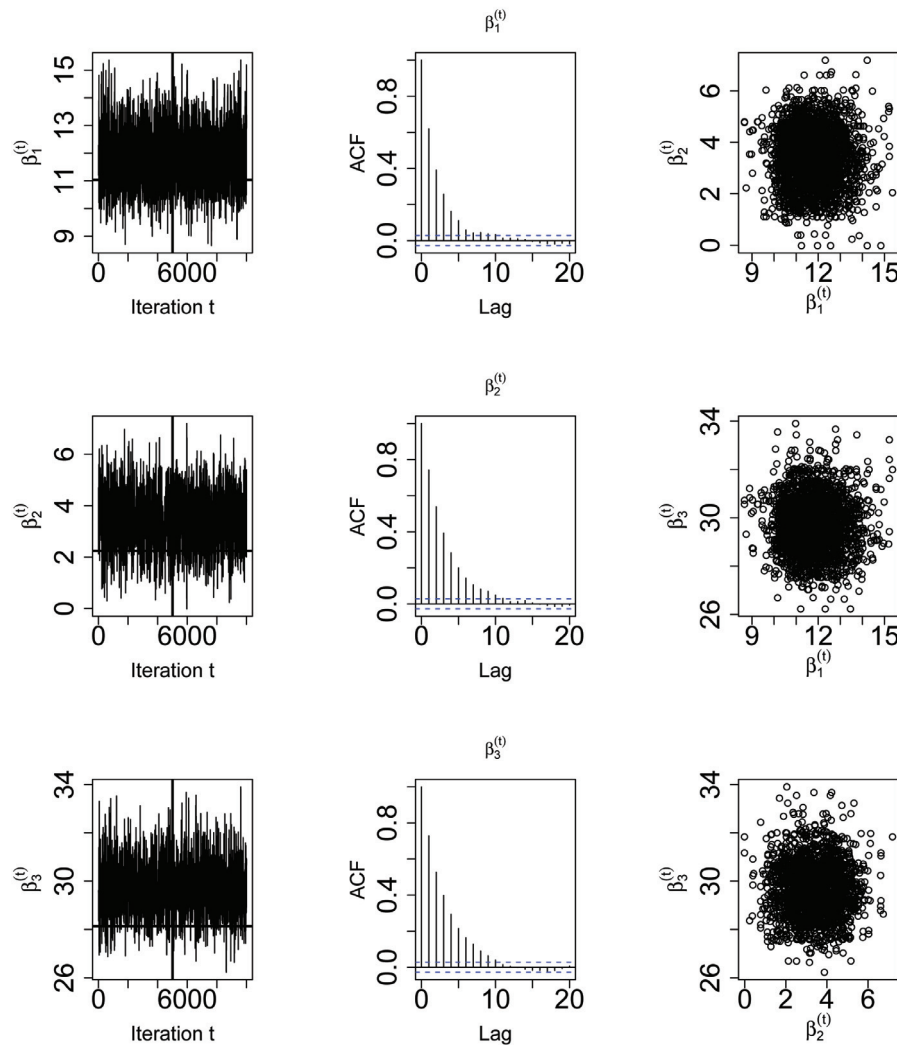


Figure 2. Results from a MCMC algorithm for sampling from the posterior $p(\boldsymbol{\beta}|r)$ where $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]$. The first column shows time series plots of sampled values $\beta_1^{(t)}$, $\beta_2^{(t)}$ and $\beta_3^{(t)}$. The vertical lines separates the burn-in phase(left) from the samples that used for future inference(right). Partial autocorrelation function are shown in the second column. Finally, $\beta_1^{(t)}$ against $\beta_2^{(t)}$, $\beta_3^{(t)}$ against $\beta_1^{(t)}$, $\beta_3^{(t)}$ against $\beta_2^{(t)}$ are shown in the third column

We provide summary statistics for the simulated posterior distributions of our models in Table 1.

Table 1. Summary of values of the mean, median, 25% and 75% quantile for two methods that we have considered

McMC method	θ_1	θ_2	θ_3
Mean	2.977	0.988	1.931
Median	2.977	0.983	1.932
25% quantile	2.847	0.945	1.892
75% quantile	3.123	1.0325	1.9708
Cholesky decomposition method	β_1	β_2	β_3
Mean	11.829	3.697	30.040
Median	11.764	3.723	30.044
25% quantile	11.183	3.057	29.339
75% quantile	12.739	4.305	30.728

To indicate the value of the transformation we provide figures for the effective sample size (Morita et al., 2008) as

well as the Iterative autocorrelation function (Sudhakar et al., 2003) and the acceptance rate for each parameter in both models in Table 2. This provides clear evidence that the transformation provides a considerable improvement in the effectiveness of the posterior simulation.

Table 2. Summary of values of effective sample size, iterative autocorrelation function and acceptance rate for the transformed proposals and the untransformed proposals

McMCMC method	θ_1	θ_2	θ_3
Effective sample size	250.76	257.95	380.91
Iterative autocorrelation function	39.87	38.77	26.25
Acceptance rate	0.0916	0.1575	0.1625
Cholesky decomposition method	β_1	β_2	β_3
Effective sample size	1568.28	1155.48	1273.411
Iterative autocorrelation function	6.376	8.654	7.852
Acceptance rate	0.37	0.230	0.233

4. Conclusion

We have shown how the a Cholesky decomposition applied to the three parameters of the Vasicek model can lead to a more efficient McMCMC estimation algorithm than simply estimating from the original formulation. It is clearly that the autocorrelation is the less than for the original parameters $\theta = [\theta_1, \theta_2, \theta_3]$ and consequently the effective sample size of the posterior simulation is considerable larger. This is a simple reparameterisation to apply, and can readily lead to better inference on the model. We believe that it would be could extend this work into the other models such as the CIR model which has a similar parameter structure.

Acknowledgements

The authors would like to thank the Editor and the reviewers for their helpful comments on an earlier version of the manuscript which have led to an improvement of this paper.

References

- Brooks, S. P., & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434-455. <http://dx.doi.org/10.1080/10618600.1998.10474787>
- Chua, C. L., Suardi, S., & Tsiaplias, S. (2013). Predicting short-term interest rates using Bayesian model averaging: Evidence from weekly and high frequency data. *International Journal of Forecasting*, 29(3), 442-455. <http://dx.doi.org/10.1016/j.ijforecast.2012.10.003>
- Cowles, M. K., & Carlin, B. P. (1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, 91(434), 883-904. <http://dx.doi.org/10.1.1.53.3445>
- Everitt, R. (2009). Using the autocorrelation time in adaptive MCMC. *Proceedings of the Sixteenth European Young Statisticians Meeting*. Retrieved from <http://www.mathos.unios.hr/eysm16>
- Gene, H., & Charles, F. (2013). *Matrix computations* (4th ed.). USA: Johns Hopkins Studies in Mathematical Sciences.
- Gilks, W. R., Richardson, S., & Siegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice* (1st ed.). London: Chapman & Hall/CRC.
- Gill, J. (2002). *Bayesian Methods: A Social and Behavioural Sciences Approach* (1st ed.). Baco Raton, BR: Chapman & Hall/CRC.
- Gray, P. (2002). Bayesian Estimation of Financial Models. *Accounting & Finance*, 42(2), 111-130. <http://dx.doi.org/10.1111/1467-629X.00070>
- Gray, P. (2005). Bayesian Estimation of Short-Rate Models. *Australian Journal of Management*, 30(1), 1-22. <http://dx.doi.org/10.1177/031289620503000102>
- Green, P. J., & Han, X. (1992). Metropolis Methods, Gaussian Proposals and Antithetic Variables. Stochastic Models, Statistical Methods, and Algorithms in Image Analysis. *Lecture Notes in Statistics*, 74, 142-164. <http://dx.doi.org/10.1007/978-1-4612-2920-9-10>

- Harris, G. (1995). Low Frequency Statistical Interest Rate Models. *5th AFIR International Colloquium*, 799-831. Retrieved from <http://www.actuaries.org/AFIR/Colloquia/Brussels/Harris.pdf>
- Iacus, S. M. (2008). *Simulation and Inference for Stochastic Differential Equations with R Examples* (1st ed.). Italy: Springer.
- Jackman, S. (2001). Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking. *Political Analysis*, 9(3), 227-241. <http://dx.doi.org/10.1093/polana/9.3.227>
- Mikosch, T. (2004). *Elementary Stochastic Calculus* (5th ed.). Denmark: World Scientific.
- Morita, S., Thall, P. F., & Müller, P. (2008). Determining the Effective Sample size of a Parametric Prior. *Biometrics*, 64(2), 595-602. <http://dx.doi.org/10.1111/j.1541-0420.2007.00888.x>
- Robert, C. P., & Casella, G. (2009). *Introducing Monte Carlo Methods with R* (1st ed.). New York, NY: Springer.
- Smith, B. J. (2007). An R Package for MCMC Output Convergence Assessment and Posterior Inference. *Journal of Statistical Software*, 21, 1-37. <http://www.jstatsoft.org/v21/i11>
- Sudhakar, R., Agarwal, R. C., & Dutta Roy, S. C. (2003). Frequency estimation based on iterative autocorrelation function. *Acoustics, Speech and Signal Processing, IEEE Transactions*, 33(1), 70-76. <http://dx.doi.org/10.1109/TASSP.1985.1164540>
- Vasicek, O. (1977). An Equilibrium Characterization of the Terms Structure. *Journal of Financial Economics*, 5(2), 177-188. [http://dx.doi.org/10.1016/0304-405X\(77\)90016-2](http://dx.doi.org/10.1016/0304-405X(77)90016-2)
- Wan, Z. (2007). *Modeling Investment Returns with A multivariate Ornstein-Uhlenbeck Process*. (Master thesis in the department of Statistics and Actuarial Science. Simon Fraser University, India). Retrieved from <http://summit.sfu.ca/item/9950/etd5855-1.pdf>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).