Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2012, **3**, 921–929 doi: 10.1111/j.2041-210X.2012.00211.x

Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation

Julien Beguin[1]*, Sara Martino[2], Håvard Rue[2] and Steven G. Cumming[1]

[1]Centre d'étude de la forêt (CEF), and Département des sciences du bois et de la forêt, Pavillon Abitibi-Price, 2405 rue de la Terrasse, Université Laval, Québec (Québec), G1V 0A6, Canada; and [2]Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

Summary

1. Spatial analysis of ecological data is central to many interesting questions in ecology. Bayesian implementation of spatially explicit models has received increasing attention from ecologists as Monte Carlo Markov Chain (MCMC) methods have become freely accessible. MCMC simulations offer a flexible framework for modelling extensive ecological data, but they also come with a wide range of problems regarding convergence, processing time and implementation.

2. We introduce to ecologists an alternative procedure for fitting Bayesian hierarchical spatial models (BHSM) with quite general spatial covariance structures. This procedure uses integrated nested Laplace approximations (INLA) as an alternative to MCMC.

3. We show, using a case study of species distribution model with binary areal data, that implementing BHSM with INLA does not require advanced programming skills, yields accurate results compared with MCMC and is rapid (e.g. a few seconds with small to moderate data sets). BHSMs efficiently removed spatial autocorrelation in the residuals and fairly evaluated uncertainty in parameter estimates and predictions.

4. The rapidity of INLA significantly decreased the processing time and allowed both sensitivity analyses on priors and cross-validation tests to be performed within a reasonable amount of time, which ultimately increased model transparency.

Key-words: Bayesian hierarchical spatial models, CAR model, Matérn, MCMC, spatial autocorrelation, species distribution model

Introduction

Spatial analysis of ecological data is central to most contemporary issues in applied ecology. Species distribution models provide key information regarding species–environment relationships, and particularly about how environmental stressors limit species distributions (Elith & Leathwick 2009). In the global context of increasing human impacts on ecosystems, predictions derived from such models can inform management plans for endangered species (Cabeza *et al.* 2004), and assessments of distributions under future land-use (Bomhard *et al.* 2005) and climate (Thuiller 2003) scenarios. Unfortunately, acquiring reliable inferences and predictions from statistical analysis of distributional data is not straightforward. One problem that must be addressed is spatial autocorrelation.

Analyses of species distributions are sensitive to spatial dependency in model residuals, or in other words, residual spatial autocorrelation (RSA; Latimer *et al.* 2006). RSA occurs when model residuals at nearby locations are not independent (Legendre 1993). RSA may arise from missing covariates that have or induce spatial structure, from incorrect specification of the functional relationship between a covariate and the response, or from neglecting to account for a spatially contagious process such as nonrandom dispersal of individuals. RSA is associated with biased Type-I error estimates owing to inflation of the effective sample size, which invalidates standard hypothesis tests (Legendre 1993). RSA may also reduce model performance (Latimer *et al.* 2006) and lead to underestimation of prediction errors (Gelfand *et al.* 2006). Hence, RSA should be avoided by correct model specification. In practice, though, important environmental covariates are often neglected because they have not been identified or are unmeasured.

Several statistical methods have emerged for modelling spatial data while accounting for RSA (e.g. see Dormann *et al.* 2007 for frequentist methods and Banerjee, Carlin & Gelfand

*Correspondence author. E-mail: julien.beguin.1@ulaval.ca

2003 for Bayesian methods). With freely available Monte Carlo Markov Chain (MCMC) methods, Bayesian implementation of spatially explicit models has received increasing attention from ecologists. The conditional autoregressive (CAR) model is one example of such models that is now routinely used in ecology for modelling spatial association in data sampled within areal units (Latimer *et al.* 2006), along roads (Thogmartin, Sauer & Knutson 2004) or transects (Aing *et al.* 2011) or at points (Haas *et al.* 2011). In the context of spatial regression, Beale *et al.* (2010) showed that Gaussian Bayesian CAR models fitted using MCMC yielded precise and unbiased parameter estimates with low Type-I and Type-II errors. While they are always possible to implement in principle, MCMC algorithms applied to complex hierarchical spatial models have a wide range of problems related to convergence and computation time: the fitting procedure is not guaranteed to converge or may converge very slowly. Moreover, implementation of the algorithms can prove problematic in itself, especially for users who are not expert in programming.

Integrated nested Laplace approximation (INLA) is a recent alternative to MCMC for fitting a large class of Bayesian models such as latent Gaussian models (Rue, Martino & Chopin 2009). Latent Gaussian models can account for hierarchical structure and non-Gaussian errors, as well as spatial and temporal autocorrelation (see INLA section). In fitting these models, INLA substitutes accurate, deterministic approximations to posterior marginal distributions in place of long MCMC simulations, thereby gaining in speed. The quality of such approximations is extremely high, as shown by comparisons with long MCMC runs (Rue, Martino & Chopin 2009). INLA has two main advantages over MCMC techniques. The first and most outstanding is computational: results can be obtained much faster than with a well-built MCMC-sampler. The INLA algorithm is naturally parallelized, thus making it possible to exploit the new computing trend of having multi-core processors. The second advantage is that INLA permits a great deal of automation and, in practice, can be almost used as a black box to analyse latent Gaussian models. This second point is especially important in applied communities where programming expertise is limited. An R-INLA library with a user-friendly R interface (R Development Core Team 2011) is freely available at http://www.r-inla.org.

Although INLA is accurate, fast and freely available, it is still little known to ecologists. The first aim of this study is therefore to introduce this recent procedure for fitting spatially explicit hierarchical models to ecological data. To do so, we first compare results obtained with a Bayesian CAR model fitted with MCMC and INLA using spatial data on the distribution of woodland caribou in eastern Canada. Second, we present and compare the performance of an alternative hierarchical spatial model using a flexible approach with a Matérn correlation function (Minasny & McBratney 2005). Using the INLA-R library, we show how these models can be easily fitted, even by nonexpert programmers. We conclude with discussion of specific issues regarding spatial autocorrelation, the interpretation of latent spatial patterns, and the parameterization of spatial random effects in hierarchical models.

Integrated nested Laplace approximation

In the following description, we assumed some familiarity with Bayesian analyses and the use of likelihood functions. For additional description of these notions, see Hilborn & Mangel (1997) or Bolker (2008).

Based upon Bayes' theorem, Bayesian analyses combine prior probability distributions with likelihood to target posterior probability distributions of parameters. In this study, we consider a particular class of Bayesian hierarchical models called latent Gaussian models. In this class of models, a latent Gaussian field x is partially observed through data y. Depending on the structure and the type of model considered, the latent field x may include, simultaneously or not, parameters associated with linear predictors (e.g. linear regression models), spline functions (e.g. additive models) and nonspatial or/and spatial random effects (e.g. hierarchical spatial models). Moreover, the probability distribution of y can follow any distribution belonging to the exponential family such as Gaussian, Poisson, Binomial, Beta and many others. The prior density of the latent field $\pi(x|\theta)$ and the likelihood of the data $\pi(y|x, \theta)$ are governed by a vector of hyperparameters θ with prior density $\pi(\theta)$. In practice, we are often interested to make inference about the posterior marginal distribution of parameters such as:

$$\pi(x, \theta|y) \propto \pi(y|x, \theta)\pi(x|\theta)\pi(\theta)$$

Getting marginal distribution of parameters, however, is almost always analytically intractable because the likelihood often is not Gaussian and the latent field is of high dimension. Stochastic sampling with MCMC methods (e.g. see Banerjee, Carlin & Gelfand 2003; Gelman & Hill 2007) can be used to solve this issue. In theory, MCMC algorithms give exact results. They also are flexible for a wide range of models. One practical limitation of MCMC for fitting spatial models, however, is its low computational speed. Processing time can be sufficiently long so as to preclude the analysis of even small spatial data sets. Moreover, computational demands can make impractical the use of sensitivity analyses on priors and cross-validation tests. This can lead to poor interpretation of the results.

As an alternative to MCMC, INLA substitutes stochastic sampling with deterministic approximations based on a clever use of the Laplace approximation and on numerical integration. INLA can be used when the latent field x is a Markov field, which means that the latent field is endowed with a conditional independent structure, for a detailed description of Gaussian Markov field see Rue & Held (2005). In addition, INLA is most useful when the main interest lies in posterior marginal distributions $\tilde{\pi}(\theta_j|y)$and $\tilde{\pi}(x_j|y)$. This is the case when, for example, assessing the effect of covariates is the goal. We give here a short description of the INLA procedure and refer the interested reader to Rue, Martino & Chopin (2009). The INLA procedure consists of three successive steps. First, an approximation of the marginal posteriors of hyperparameters $\tilde{\pi}(\theta|y)$ is computed using Laplace approximation. The main use of this approximation is to select good estimation points to integrate out the uncertainty with respect to θ when approximating the posterior marginals of each parameter x_i. A key feature here is to avoid representing $\tilde{\pi}(\theta|y)$ parametrically, which allows more flexibility together with reducing computational demands. The second step consists of approximating the posterior marginal for the x_i conditioned on selected values of θ, $\tilde{\pi}(x_i|\theta, y)$, using again Laplace approximations. Finally, an approximation to $\tilde{\pi}(x_j|y)$ is computed via numerical integration as:

$$\tilde{\pi}(x_j|y) = \int \tilde{\pi}(x_i|\theta, y)\tilde{\pi}(\theta|y)\mathrm{d}\theta = \sum_j \tilde{\pi}(x_i|\theta_j, y)\tilde{\pi}(\theta_j|y)\Delta_j \qquad \text{eqn 1}$$

Where the points θ_j are selected in the first step of the INLA algorithm and Δ_j are associated weights. When the number of hyperparameters is small (e.g. < 6), as is the case in most ecological applications, the computation of these approximations is very fast and accurate, as shown by long-run comparisons with MCMC (see Rue, Martino & Chopin 2009). These computational advantages and the fact that there is a user-friendly software make INLA an appealing alternative for latent Gaussian models, including generalized linear (or additive) mixed models, time-series models, geoadditive models and state-space models, all of which are currently applied more or less commonly to the analysis of ecological data.

Case study: Woodland caribou in the boreal forest of eastern Canada

We illustrate the use of INLA in the context of hierarchical spatial modelling with spatially autocorrelated distributional data. First, we compared the performance of INLA with MCMC using an intrinsic spatial CAR model. For the MCMC method, we used a Gibbs sampler that was implemented in WinBUGS (available online at http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml). Second, we show an alternative hierarchical spatial method, which uses a Matérn correlation function to account for spatial associations among units. Matérn spatial correlation models are not supported in WinBugs, so we only present these result using INLA. Overall, the purpose of these comparisons is to show that powerful alternatives to MCMC exist for fitting a wide range of spatially explicit hierarchical models in ecology and evolution. All models using INLA were fitted in R v2.14.1 using the R-INLA library. Data and R-codes are available in Appendix 1.

DATA

Our case study consists of species distribution data obtained in 1999 from an extensive aerial winter survey of forest-dwelling caribou (*Rangifer tarandus caribou*) in a region of 42 071 km^2 located in the boreal forest of eastern Québec, Canada (Courtois *et al.* 2003). This region and data are a subset of those studied by Fortin *et al.* (2008). The observational units for this study are the presences or absences of intensive caribou snow track networks (ICTN). An ICTN indicates areas that were used briefly and intensively by a small group of caribou for foraging or shelter. A total of 232 ICTNs were detected and mapped by an exhaustive winter aerial survey consisting of a dense grid of fixed-wing transects backed up by helicopter spot-checks (see Courtois *et al.* 2003 for details). The mean size of an ICTN is 0·53 km^2 (SE, ± 0·06 km^2). ICTN locations were taken to be their centroids. The probability of detection given presence is high (> 90%), owing to strong contrasts between ICTN and adjacent undisturbed snow surfaces (Courtois *et al.* 2003). Given the high detectability of ICTN and the intensity of sampling effort, false-positives and false-negatives can be ignored. We initially selected 13 covariates that could potentially influence the probability of ICTN presence (Table 1). These covariates have been obtained from several data sources, including digital forest inventories, interpolated climate data (McKenney *et al.* 2006) and a 0·75 arc-second resolution Digital Elevation Model (available online at http://www.geobase.ca/geobase/fr/data/cded/description.html). The spatial resolution of covariates was ~500 m^2 for elevation, 8 ha for forest inventory attributes and ~100 km^2 for climate data. All attributes were calculated as total areas or averages over a common 100 km^2 grid defined by the climate data (see Table 1). A linear model of coregionalization revealed that a grain size of 100 km^2 was appropriate to delineate homogeneous spatial units within the study region. The chosen spatial grid and grain size also corresponds to that of our interpolated climate data, the covariate of lowest spatial resolution. Hereafter, each elementary unit of this grid is referred to as a cell ($N = 465$). Caribou presence/absence data within cells, however, was calculated for a 1-km^2 resolution subgrid. This resolution corre-

Table 1. Environmental variables used as predictors in our study. Minimum, mean and maximum values of each explanatory variable, at the scale of 100 km^2 grid cells, are presented for the entire region

Environmental variables	Minimum	Mean	Maximum
Wetlands (km^2)	0·1	2·8	23·1
Water bodies (km^2)	1·9	14·0	87·8
Lichen woodland (km^2)*	0	3·6	27·8
Bare land dominated with lichen (km^2)†	0	3·5	32·2
Coniferous Forests (km^2)	0	54·1	91·5
Deciduous Forests (km^2)	0	3·5	44·7
Wildfire (km^2)	0	9·8	82·2
Logging (km^2)	0	19·9	83·1
Other land cover (km^2)	0	3·1	43·4
Road density (km/km^2)	0	0·5	5·1
Winter mean temperature (°C)	−18·91	−16·5	−12·67
Winter total precipitation (mm)	47	60	73
Elevation (m)	239	500	879

Variables in bold type were retained in the final candidate model for each statistical method.

* or WoodLichen in Fig. 3.

† or OpenLichen in Fig. 3.

sponds to roughly twice the mean area of an ICTN (see Fig. 1).

STATISTICAL MODELS

Let n_i be the number of 1-km^2 subcells within the ith cell, and Y_i the number of these subcells where the presence of ICTN is observed. We assume Y_i to be a binomial random variable and have modelled the probability p_i of ICTN presence within cells using a logit link function and the following generic model:

$$\begin{aligned} Y &\sim \text{Binomial}(n_i, p_i) \\ \text{logit}(p_i) &= \boldsymbol{\beta}\mathbf{X}_i + f(s_i) \end{aligned} \qquad \text{eqn 2}$$

Here, $\mathbf{X}_i$ is the vector of covariates for cell i, and $\boldsymbol{\beta}$ is the vector of parameters to be estimated. The two hierarchical spatial models considered in this study only differ in their intrinsic way of defining the spatial random effect $f(s)$. A spatial random effect is used to model dependence among neighbouring cells that is not explained by the covariates.

All models assumed a vague Gaussian prior for the regression parameters $\boldsymbol{\beta} \overset{iid}{\sim} N(\text{mean} = 0, \text{precision} = 0{\cdot}001)$, where precision = 1/variance. For the binomial intrinsic CAR model fitted with INLA, the prior for the spatial random effect is defined conditionally as:

$$f_s(s)|f_s(s'), s \neq s', \lambda_s \sim N\left(\frac{1}{n_s}\sum_{s\sim s'} f_s(s'), \frac{1}{n_s\lambda_s}\right)$$

where $s{\sim}s'$ indicates that the two cells s and s' are neighbours and n_s is the number of neighbours of cell s. We defined two cells to be neighbours if they directly share a single boundary point. The unknown precision hyperparameter λ_s controls the smoothness of the spatial random effect and the prior is defined on a logarithmic scale such as it follows a logGamma (shape = 24·47, scale = 0·001) distribution (see the following section on prior choices). Further details of this model can be found in chapter 3 in Rue & Held (2005).

The spatial random effect in the Matérn model is a Markov representation on a regular grid of a continuous Gaussian field with a Matérn covariance function. The Matérn function offers a flexible way of modelling spatial dependence among units as it includes the exponential and squared exponential covariance functions as special cases (see section 2.1.3. in Banerjee, Carlin & Gelfand 2003 or Minasny & McBratney 2005 for details).

The correlation function in the Matérn model is defined as:

$$\text{Corr}(d) \propto (kd)^{v} K_v(kd)$$

Where d is the Euclidean distance, $K_v()$ is the Bessel function of order v. The range is defined to be $\sqrt{8}/k$and it is the distance at which two cells are practically uncorrelated. The Matérn model has two hyperparameters: the precision of the spatial random effect and the range as defined above. As it is a difficult task to estimate simultaneously both range and precision parameters, we fixed the range to be five times the distance between two neighbouring cells. This value for the range was meant to capture the fine scale spatial association among cells. As for the CAR model, the precision hyperparameter is defined internally on a logarithmic scale and we assigned it a logGamma (shape = 23·36, scale = 0·001) prior. The Matern covariance function implies a dense covariance matrix that greatly increases computational demands and processing time. The INLA software uses a Markov representation of the Matern field that has been introduced by Lindgren, Lindstrøm & Rue (2011). The Markov representation offers several com-

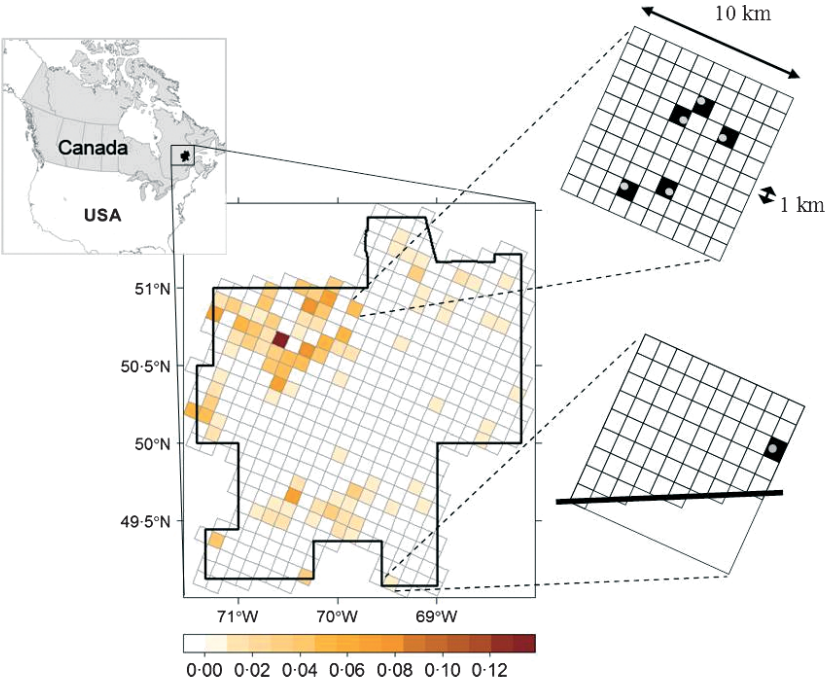

Fig. 1. Study region located in the boreal forest of eastern Canada. Bold black lines in the middle-panel delineate boundaries of the aerial survey. Colours define the proportion of intensive winter track networks of woodland caribou (ICTN) in each of 465 cells of 100-km^2. The right-side panel highlights two spatial unit structures. The 100-km^2 grid represents spatial units at which environmental variables are evaluated, and the 1-km^2 grid represents spatial units where the presence (black squares, with grey ICTN centroids inside) or absence (white squares) of ICTN was detected. See text for further description.

putational advantages and greatly reduces the running time to fit such model. The description of such Markov representation is beyond the scope of this study and the interested reader is referred to Lindgren, Lindstrøm & Rue (2011).

Finally, to ensure parameter identifiability, a sum-to-zero constraint is imposed on *f*(*s*) in all cases.

VARIABLE SELECTION AND PRIOR CHOICES

Logistic, CAR and Matérn models were fit using the same design matrix $\mathbf{X}_i$ of environmental covariates, which were selected from the initial set of 13 covariates by a preliminary screening process. Selection was carried out by fitting a full logistic model without spatial random effect and then calculating variance inflation factors (VIF) as a test for multicollinearity (Zuur *et al.* 2009). Covariates with VIF $>$ 4 were removed one-by-one using a backwards, stepwise procedure. We then fit every model with the same eight covariates using linear effects on the logit scale (Table 1). For the CAR and Matérn models, the choice of the shape parameter for the precision of the spatially structured effect *f*(*s*) determines the smoothness of the spatial effect and, through this, the spatial scale at which it operates. Misspecification of this parameter can lead to poor interpretation of the results (e.g. overfitting), so special care must be taken. To avoid pitfalls, we manually investigated the sensitivity of the results for a wide range of shape parameter values. In particular, we investigated the effect of various shape parameter values on coefficient estimates and credible intervals of explanatory variables, deviance information criterion (DIC), the number of effective parameters and the number of effective replicates. As model comparisons that are based on DIC can underpenalize for model complexity and thus encourage overfitting (Plummer 2008), we retained values of the shape parameters such that the ratio of data/parameters was always $>$20 (Burnham & Anderson, 2002).

MODEL OUTPUTS

For each fitted model, we calculated the mean predicted value of ICTN presence in each cell (with 95% credible intervals) and estimated RSA using variograms. We compared inferential properties among models using parameter estimates and their credible limits. We assessed model goodness-of-fit by calculating the Pearson's correlation coefficient (*r*) between predicted values and observed proportions of ICTN presence at the cell level. Finally, we used '*leave-one-out*' cross-validation to assess the predictive power of each model, based on the conditional predictive ordinate (CPO) statistic (Held, Schrödle & Rue 2010). Following Held, Schrödle & Rue (2010), the CPO value for the *i*th cell is defined as $\text{CPO}_i = \pi(y_{i,\text{obs}}|y_{\text{-}(i),\text{obs}})$, where $y_{i,\text{obs}}$ is the binomial outcome of the ICTN presence for cell *i*, and $y_{\text{-}(i),\text{obs}}$ denotes the data without the *i*th cell. The number of cells is 465. We used the mean logarithmic-CPO_i defined as:

$$\text{LCPO} = -\frac{1}{465}\sum_i \log(\text{CPO}_i) \qquad \text{eqn 3}$$

The lower the LCPO for a model, the better is its predictive power.

Results

Conditional autoregressive and Matérn models adequately removed RSA present in the nonspatial logistic model (Fig. 2). We detect no difference in parameter estimates for the Bayesian CAR model using MCMC vs. INLA approaches (Fig. 3). Nor did we find any significant difference in parameter estimates between Matérn and CAR models fitted with INLA (Fig. 3). When fitted with MCMC (100 000 iterations, three chains), the CAR model took several hours to converge, whereas INLA took *c.* 5 s. Estimation of the Matérn model with INLA was also a matter of seconds.

As expected, results of hierarchical spatial model outcomes are sensitive to specification of the shape parameter (Fig. 4, Appendix 2). The shape parameter governs the smoothness of the spatial random effect: low values (e.g. uninformative priors) mirror spatial associations at short distances, whereas high values mirror spatial correlation at large distances (Fig. 4). Our result shows that accounting for autocorrelation at short distances comes with the cost of increasing number of estimated parameters, which in turn increases the risk of overfitting (Fig. 4). We found that DIC-based comparisons underpenalized for model complexity and encouraged overfitting (Fig. 4). To avoid this issue and to balance the amount of information contained in the data with the number of estimated parameters, we selected a shape parameter so that the

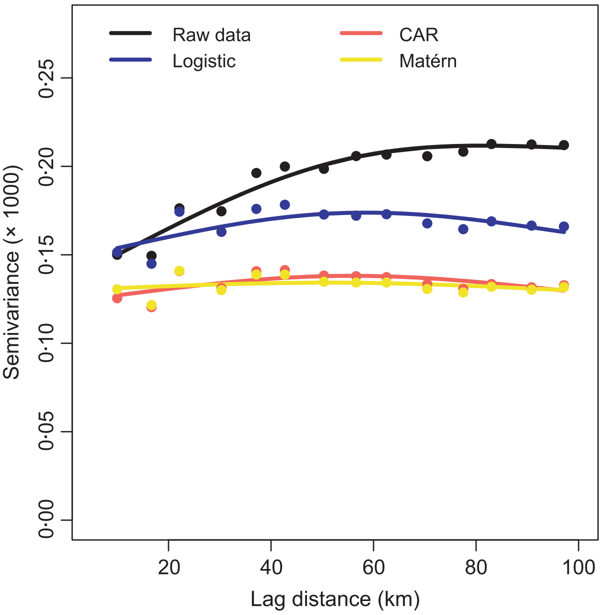

Fig. 2. Spatial variograms of (i) the raw data; (ii) residuals of the non-spatial logistic model; (iii) residuals of the hierarchical conditional autoregressive (CAR) model; and (iv) residuals of the spatial model with a Matérn correlation function. Note: A likelihood-ratio (LR) test between the following two models shows that residuals of logistic model are spatially autocorrelated (LR = 10·1, *P* = 0·007): model 1: Semi-variance = β_0. model 2: Semi-variance = $\beta_0 + \beta_1$ distance + β_2 distance^2.

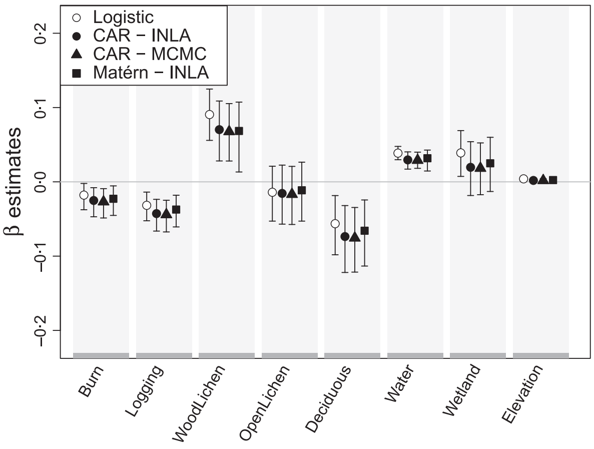

Fig. 3. Parameter estimates (±95% credible intervals) of every environmental variable used for each statistical method in this study. See Table 1 for a description of each variable.

number of effective replicates was greater than, but close to, 20 (Fig. 4). Lower values of effective replicates encouraged overfitting, while higher values only mirrored spatial trend at broad scales (Fig. 4). When values of the shape parameter are ≥24·5, the binomial CAR model no longer differs from a nonspatial logistic model.

The occurrence probability of track networks of woodland caribou increased with the proportion of lichen woodland and water bodies in a cell, but decreased with the portion of deciduous stands and disturbances caused by fire and logging (Fig. 3). The effect of logging was consistently more negative than that of burning. No significant selection pattern was found for bare lands dominated by terrestrial lichens or for wetlands (Fig. 3). The occurrence probability of track networks tended to increase with elevation, but elevation also showed a greater sensitivity to specification of the shape parameter (Appendix 2).

The distribution of spatial random effects among Bayesian models revealed strong spatial patterns at multiple scales (Figs 4 and 5). At coarse scales, favourable winter areas for woodland caribou increased from northeast to west. At intermediate scales, two suitable subregions emerged in the south and northwest parts of the study area. Within these subregions, finer-scale local variation was evident. Although CAR and

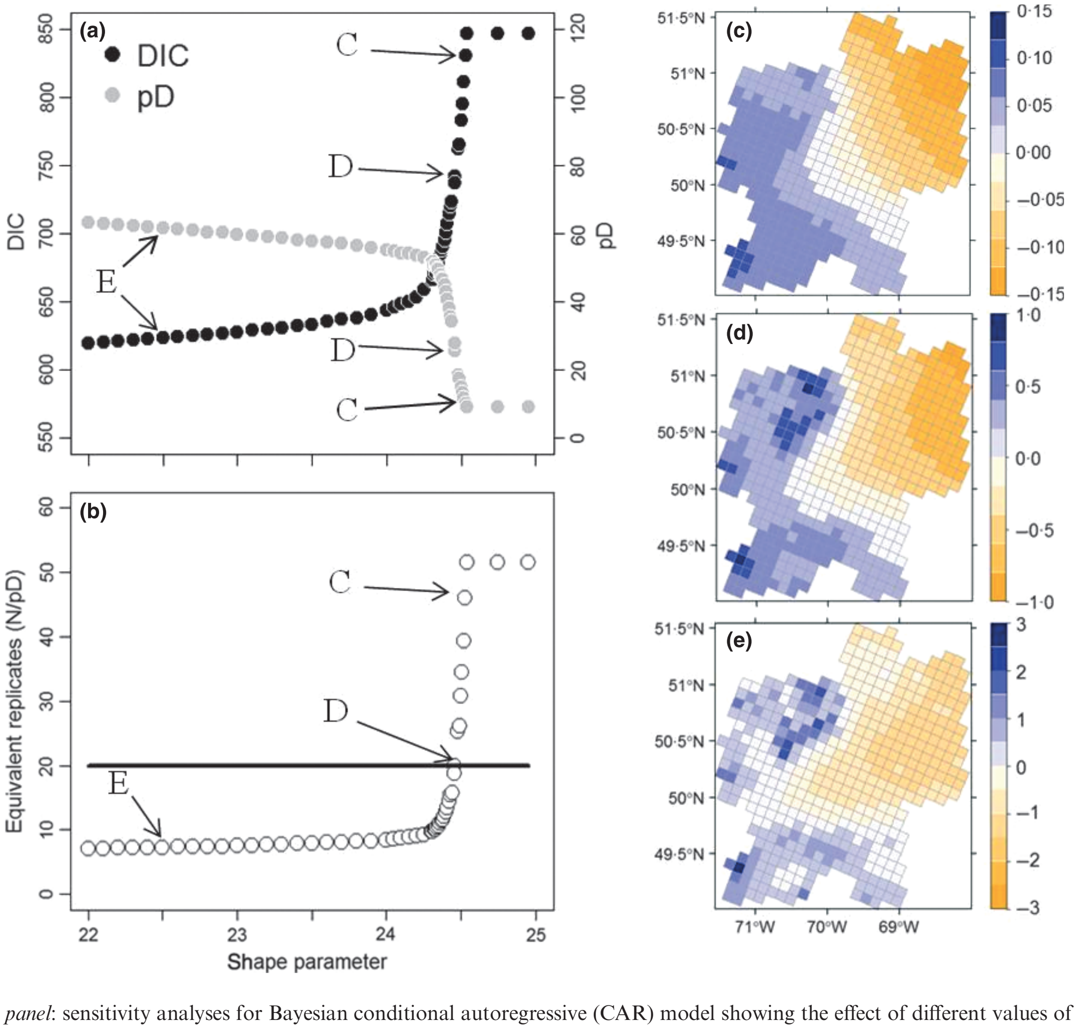

Fig. 4. *Left panel*: sensitivity analyses for Bayesian conditional autoregressive (CAR) model showing the effect of different values of the shape parameter on (a) deviance information criteria (DIC) and the number of effective parameters (pD), and (b) the number of effective replicates. Letters C, D and E correspond to shape parameters of 24·52, 24·47 and 22·50, respectively. *Right panel*: posterior mean of the spatial random effect of Bayesian CAR model for each shape parameter defined above.

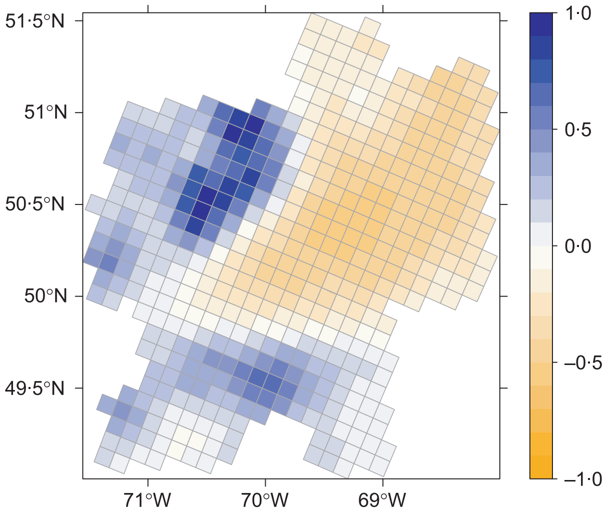

Fig. 5. Posterior mean of the spatial random effect for hierarchical spatial model with Matérn correlation function.

Matérn models roughly had the same number of effective parameters, the Matérn model allowed describing a latent spatial structure at finer scale, likely because of the flexibility of the Matérn correlation function and because data are on a regular lattice.

Overall, CAR and Matérn models improved predictive power by 10% and the percentage of explained deviance by 14% compared with nonspatial logistic model (Fig. 6). Uncertainty of the estimates, however, is 5–20% larger for the Bayesian than for the nonspatial logistic model (Fig. 3). As expected, models showing higher uncertainty in parameter estimates had higher prediction errors (Fig. 6). In no case, was a parameter estimate significant in a spatial and nonspatial model but of differing sign (Fig. 3).

Discussion

Making reliable inferences and robust predictions from the analysis of spatial data is central to many pressing issues in ecology and evolution. Hierarchical spatial models are now routinely used to analyse spatially autocorrelated data as they offer convenient properties such as a flexibility regarding the probability distribution of the response variable, together with their ability to evaluate simultaneously the contributions of fixed predictors and spatial random effects to the likelihood. This last property ensures that variability is properly attached to predictions (Gelfand *et al.* 2006) and, if needed, allows predictions to be made at unsampled locations, thereby accommodating gaps in sampling and irregular sampling intensity. In this study, we evaluated a relatively new statistical methodology for estimating hierarchical spatial models, which uses the INLA (Rue, Martino & Chopin 2009) as an alternative to MCMC simulations. With this novel numerical inference approach, MCMC sampling becomes redundant as the posterior marginal distributions are accurately approximated in a fully automated fashion (Held, Schrödle & Rue 2010).

Several remarks regarding INLA can be drawn from our study. First, the results of CAR model that were obtained with INLA were so accurate that no tangible difference in model outcomes could be detected with MCMC. This confirms previous comparative results obtained for a wide range of Gaussian latent models (Rue, Martino & Chopin 2009). Second, the processing time for fitting hierarchical spatial models was very rapid, which ultimately allowed us to perform cross-validation

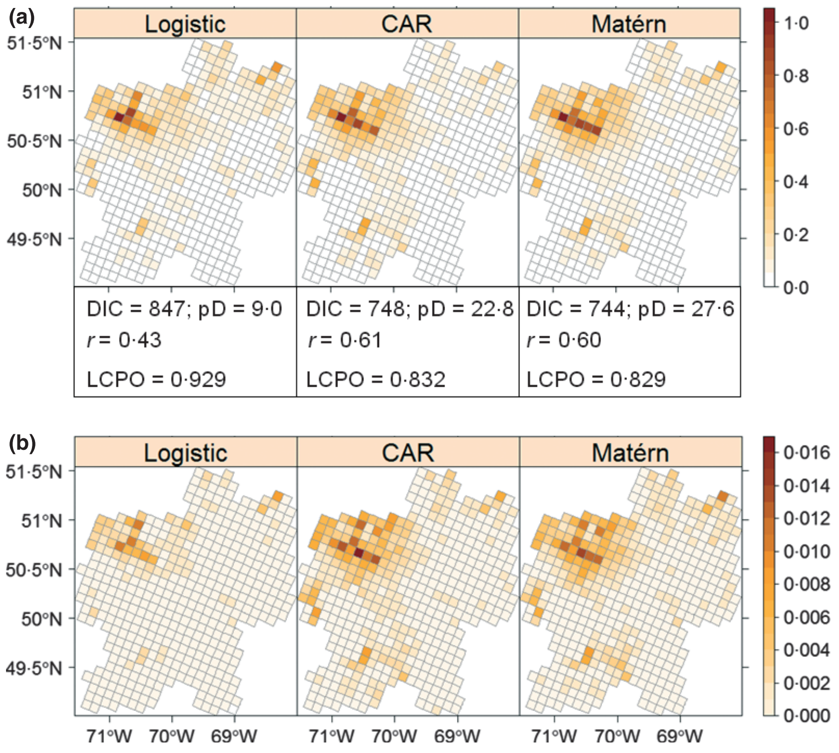

Fig. 6. (a) Mean predicted probability of occurrence of woodland caribou track networks (values have been rescaled to range from 0 to 1); and (b) Standard deviation of the predicted probability of occurrence for the statistical methods tested in this study. Below panel (a), a summary table shows for each statistical method: (1) Deviance information criteria (DIC); (2) the effective number of parameters (pD); (3) the correlation coefficient between observed and predicted values (*r*); (4) the log-score statistic (LCPO) measuring the predictive power of each model through cross-validation (the lower, the better; see text for further description).

tests and sensitivity analyses on priors within a reasonable amount of time. Such a gain in processing time with INLA opens new perspectives in modelling spatial data under the Bayesian framework, as it alleviates one of the most important bottlenecks associated with MCMC. Third, the use of INLA with the R interface greatly facilitates implementation of Bayesian hierarchical spatial models. A strong expertise in programming is thus no longer an obstacle to fitting these models. Fourth, an increasing variety of hierarchical spatial models are now available for end-users (e.g. spatial GLMM, geographic weighted regression, thin-plate splines, among others), which ultimately broadens the range of tools available to ecologists. These models, together with the models presented in this study, allow a wider range of questions to be addressed in ecology and evolution. As with any statistical procedure, INLA also has limitations: precision of approximations becomes less accurate as the number of hyperparameters to be estimated increases above 6; occupancy models containing an observational process to cope with imperfect detection (MacKenzie *et al.* 2002) are not yet implemented in INLA and will require further developments. Site-occupancy models developed for spatially correlated observations in an active area of research (Aing *et al.* 2011); although INLA greatly reduces the processing time relative to MCMC on small to moderate-sized data sets, the spatial modelling of large data sets ($N > 10^6$) remains challenging.

In hierarchical spatial models, the shape parameter of the gamma distribution that was used as a prior for the precision of the spatial model controls the smoothness of the spatial random effect. The degree of smoothness can be viewed as the spatial equivalent of the pooling factor ω (Gelman & Hill 2007), which represents the degree to which the estimates are pooled together rather than estimated separately for each group factor. Our case study emphasized that parameterization of the shape parameter requires special care. The choice of an uninformative prior favours a low degree of smoothing, which in turn can yield a model that is overfitted to the given observations (see Fig. 4a). As described by Plummer (2008), this situation typically occurs when the choice of shape parameter is solely based on DIC comparisons, as DIC can underpenalize model complexity in this class of models. In contrast, highly informative priors increase the degree of spatial smoothing to such a point that the full model becomes equivalent to a nonspatial one (see Fig. 4a, b). Up to now, there is no Bayesian equivalent to AICc (see Burnham & Anderson 2002). Hence, we recommend that, based on information theory, the value of the shape parameter be defined so that the ratio between sample size and the number of effective parameters be >20. This stopping rule maintains a balance between information contained in the data and the number of effective parameters (Rue, Martino & Chopin 2009). Sensitivity analyses allow such thresholds to be identified and should be an integral part of any modelling exercise with hierarchical models. Additional work is needed to define Bayesian comparative criteria sensitive to small numbers of equivalent replicates.

Our results on landscape selection patterns largely support previous findings that show woodland caribou avoid regeneration areas following recent disturbance (Vors *et al.* 2007; Courbin *et al.* 2009) and deciduous stands (Courbin *et al.* 2009), which likely mimics a behavioural response to increasing predation risk (Wittmer *et al.* 2007). For example, deciduous stands and regenerating areas are used by moose (*Alces alces* L.), and increasing moose abundance often translates into higher risk of predation by wolves (Seip 1992; Wittmer *et al.* 2007). The positive selection of water bodies presumably reflects a similar strategy as, on one hand, these large ice-covered open areas improve the ability to detect predators and, on the other hand, surrounding forests enhance the ability to escape from predators (Mysterud & Ostbye 1999). We found a strong positive selection for lichen woodland that confirms the importance of this habitat in providing arboreal and terrestrial lichens, two major components of woodland caribou winter diet (Johnson, Parker & Heard 2001). Our results also show that, in addition to removing RSA, the inclusion of a spatial random effect yields strong latent spatial patterns at multiple scales that were not explained by environmental covariates. We argue that these latent spatial patterns offer opportunities for researchers to investigate and make further ecological hypotheses about the underlying processes that generated these patterns. For instance, the scale and distribution of spatial random effects could be used as a surrogate in inferring processes from spatial patterns (McIntire & Fajardo 2009). These effects might also be indicative of important missing covariates and further serve sampling strategies by prioritizing locations where data acquisition seems to be the most urgent.

Acknowledgements

We are grateful to Daniel Fortin, Réhaume Courtois, Claude Dussault and the Ministère des Ressources naturelles et de la Faune du Québec for providing access to the caribou data, and the Géo/Stat Centre at Laval University for the forest inventory data. We also thank D. McKenney and P. Papadopol of the Canadian Forest Service, Natural Resources Canada, for providing climate data over the study region. Funding for this research was provided by an NSERC (Natural Sciences and Engineering Research Council of Canada) scholarship to J.B. We sincerely thank P. Racine for his contributions in manipulating GIS data, E.J.B. McIntire and G. Daigle for stimulating discussions on spatial analyses, W.F.J. Parsons for English revisions, and the associate editor together with three anonymous reviewers for their constructive comments.

References

Aing, C., Halls, S., Oken, K., Dobrow, R. & Fieberg, J. (2011) A Bayesian hierarchical occupancy model for track surveys conducted in a series of linear, spatially correlated, sites. *Journal of Applied Ecology*, **48**, 1508–1517.

Banerjee, S., Carlin, B.P. & Gelfand, A. (2003) *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall, Boca Raton, Florida, USA.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.

Bolker, B.M. (2008) *Ecological Models and Data in R*. Princeton University Press, Princeton, New Jersey, USA.

Bomhard, B., Richardson, D.M., Donaldson, J.S., Hughes, G.O., Midgley, G.F., Raimondo, D.C., Rebelo, A.G., Rouget, M. & Thuiller, W. (2005) Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. *Global Change Biology*, **11**, 1452–1468.

Burnham, K. P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, USA.

Cabeza, M., Araújo, M.B., Wilson, R.J., Thomas, C.D., Cowley, M.J.R. & Moilanen, A. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology*, **41**, 252–262.

Courbin, N., Fortin, D., Dussault, C. & Courtois, R. (2009) Landscape management for woodland caribou: the protection of forest blocks influences wolf-caribou co-occurrence. *Landscape Ecology*, **24**, 1375–1388.

Courtois, R., Gingras, A., Dussault, C., Breton, L. & Ouellet, J.P. (2003) An aerial survey technique for the forest-dwelling ecotype of Woodland Caribou, *Rangifer tarandus caribou*. *Canadian Field Naturalist*, **117**, 546–554.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kuhn, I., Ohlemuller, R., Peres-Neto, P.R., Reineking, B., Schroder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, **40**, 677–697.

Fortin, D., Courtois, R., Etcheverry, P., Dussault, C. & Gingras, A. (2008) Winter selection of landscapes by woodland caribou: behavioural response to geographical gradients in habitat attributes. *Journal of Applied Ecology*, **45**, 1392–1400.

Gelfand, A.E., Silander, J.A., Wu, S.S., Latimer, A., Lewis, P.O., Rebelo, A.G. & Holder, M. (2006) Explaining species distribution patterns through hierarchical modeling. *Bayesian Analysis*, **1**, 41–91.

Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York, New York, USA.

Haas, S.E., Hooten, M.B., Rizzo, D.M. & Meentemeyer, R.K. (2011) Forest species diversity reduces disease risk in a generalist plant pathogen invasion. *Ecology Letters*, **14**, 1108–1116.

Held, L., Schrödle, B. & Rue, H. (2010) *Posterior and Cross-validatory Predictive Checks: A Comparison of MCMC and INLA in Statistical Modelling and Regression Structures – Festschrift in Honour of Ludwig Fahrmeir*. Springer Verlag, Berlin, Germany.

Hilborn, R. & Mangel, M. (1997) *The Ecological Detective: Confronting Models with Data*. Princeton University Press, Princeton, New Jersey, USA.

Johnson, C., Parker, K. & Heard, D. (2001) Foraging across a variable landscape: behavioral decisions made by woodland caribou at multiple spatial scales. *Oecologia*, **127**, 590–602.

Latimer, A.M., Wu, S.S., Gelfand, A.E. & Silander, J.A. (2006) Building statistical models to analyze species distributions. *Ecological Applications*, **16**, 33–50.

Legendre, P. (1993) Spatial autocorrelation – trouble or new paradigm? *Ecology*, **74**, 1659–1673.

Lindgren, F., Lindstrøm, J. & Rue, H. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, **73**, 423–498.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.

McIntire, E.J.B. & Fajardo, A. (2009) Beyond description: the active and effective way to infer processes from spatial patterns. *Ecology*, **90**, 46–56.

McKenney, D.W., Pedlar, J.H., Papadopol, P. & Hutchinson, M.F. (2006) The development of 1901–2000 historical monthly climate models for Canada and the United States. *Agricultural and Forest Meteorology*, **138**, 69–81.

Minasny, B. & McBratney, A.B. (2005) The Matérn function as a general model for soil variograms. *Geoderma*, **128**, 192–207.

Mysterud, A. & Ostbye, E. (1999) Cover as a habitat element for temperate ungulates: effects on habitat selection and demography. *Wildlife Society Bulletin*, **27**, 385–394.

Plummer, M. (2008) Penalized loss functions for Bayesian model comparison. *Biostatistics*, **9**, 523–539.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rue, H. & Held, L. (2005) *Gaussian Markov Random Fields: Theory and Applications*. Volume 104 of Monographs on Statistics and Applied Probability. Chapman and Hall, London.

Rue, H., Martino, S. & Chopin, N. (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, **71**, 319–392.

Seip, D.R. (1992) Factors limiting woodland caribou populations and their interrelationships with wolves and moose in southeastern British Columbia. *Canadian Journal of Zoology*, **70**, 1494–1503.

Thogmartin, W.E., Sauer, J.R. & Knutson, M.G. (2004) A hierarchical spatial model of avian abundance with application to cerulean warblers. *Ecological Applications*, **14**, 1766–1779.

Thuiller, W. (2003) BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.

Vors, L.S., Schaefer, J.A., Pond, B.A., Rodgers, A.R. & Patterson, B.R. (2007) Woodland caribou extirpation and anthropogenic landscape disturbance in Ontario. *The Journal of Wildlife Management*, **71**, 1249–1256.

Wittmer, H.U., McLellan, B.N., Serrouya, R. & Apps, C.D. (2007) Changes in landscape composition influence the decline of a threatened woodland caribou population. *Journal of Animal Ecology*, **76**, 568–579.

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A. & Smith, G.M. (2009) *Mixed Effects Models and Extensions in Ecology with R*. Springer Science, New York, USA.

Received 30 January 2012; accepted 2 April 2012
Handling Editor: Matthew Spencer

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. R-code for running with INLA all hierarchical models presented in the paper.

Appendix S2. Effect of shape parameter on coefficients of regression (±95% CI).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be reorganised for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.