

Disentangling spatio-temporal processes in a hierarchical system: a case study in fisheries discards

Mafalda Viana, Andrew L. Jackson, Norman Graham and Andrew C. Parnell

M. Viana (vianam@tcd.ie) and A. L. Jackson, Dept of Zoology, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland. MV and ALJ also at: Centre for Biodiversity Research, Trinity College Dublin, Dublin 2, Ireland. – N. Graham, Fisheries Sciences Services, Marine Inst., Rinville, Oranmore, Co. Galway, Ireland. – A. C. Parnell, School of Mathematical Sciences (Statistics), Complex and Adaptive Systems Laboratory, Univ. College Dublin, Dublin 4, Ireland.

In the last decade, various spatial and temporal methodologies were developed to investigate the processes that drive ecological and evolutionary patterns. However, these methods frequently fail to acknowledge that the observed patterns result from the overlap of different underlying processes. In order to understand how the patterns are formed, we must have recourse to methods that allow us to disentangle these simultaneous processes. Here we develop a hierarchical spatial predictive process (PP) combined with a separable temporal PP to disentangle and describe those overlapping processes in one very frequent setting in ecology and evolution: multilevel spatio-temporally indexed data. We present our methodology through a case study of fisheries discards and investigate for example whether the inclusion of the hierarchical structure and the temporal processes of the system alter the observed spatial patterns. Recently it is recognized that understanding the processes driving discards is essential to sustainably manage and conserve marine resources. The results show that consideration of multiple underlying processes dramatically changes the pattern and characteristics of the discards hot- and coldspots. In the Irish Sea, the inclusion of the hierarchical structure of the system leads to the reduction of the hot- and coldspots. Simultaneously, our model identifies key bi-annual fluctuations in the temporal process which, together with the variance associated at the level of individual fishing trips in the hierarchical structure of the data explained most of the variance driving discards. Whether the hierarchical, spatial and temporal processes are considered together or not can profoundly alter our understanding of what constitutes an appropriate mitigation measure. Misidentification of hotspots can culminate in inappropriate mitigation practices which can sometimes be irreversible. As the proposed method offers a unified approach for understanding the processes that drive observed patterns, many areas in ecology such as conservation and epidemiological studies can benefit from its use, increasing the effectiveness of management plans.

Elucidating the links between emergent patterns and the functions and processes driving them has been a central challenge in ecology (MacArthur 1984, Levin 1992, Gustafson 1998, Urban 2006). However, for statistical simplicity we often ignore that the observed patterns result from the overlap of different mechanistic processes. This is particularly relevant in fields such as spatial ecology, conservation and evolutionary studies in which we frequently struggle with the confounding effects of space, time and the structure in which the system is embedded and constrained (e.g. environment or genes).

For example, it is common for two biological units (those being cells, individuals or populations) geographically close to each other to be more similar or have a stronger relationship than those far apart (e.g. spatial autocorrelation, Fig. 1a) (Legendre et al. 2002, Dormann et al. 2007); additionally, these biological units are frequently organized in groups that are further divided in subgroups and so forth over multiple

scales (multilevel structure, Snijders and Bosker 1999) which can create additional correlation beyond that imposed by the spatial autocorrelations (Fig. 1b). These groups can arise from the sampling design or as a result of naturally occurring functional units such as a genealogical/phylogenetic tree, ecological niches within food webs or a network of connected lakes. Finally, the relationships between biological units can change over time according to a temporal process in which successive biological units are expected to be more similar than those distant in time, or simply vary with factors associated with time (e.g. temporal autocorrelation or seasonal effects owing to environmental conditions). To fully understand the pattern generated by these biological units we must therefore distinguish the effect of these multiple overlapping processes.

In the last decade, various methodologies were developed to independently investigate these effects, e.g. GAMs, and Kriging (Cressie 1993) for spatial patterns, hierarchical and

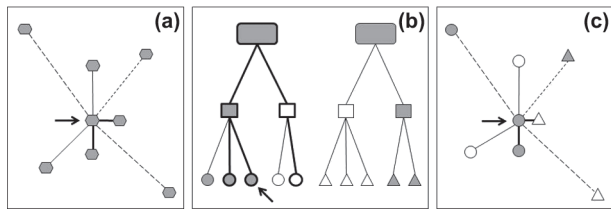


Figure 1. (a) Spatial autocorrelation diagram. The focal datum indicated by the arrow is more correlated to data close in space than those further away as represented by the line thickness. (b) Multi-level structure diagram. The focal datum, located at the lowest level of the structure, may be most similar to other datum from within the same medium level entity (filled circles within squares), but are also correlated with datum from other entities (open circles) within the same upper level. (c) Joint spatial and multilevel correlations diagram. The focal datum is both close in space to data derived from the same level (closed circles) or the level above (open circles) but also from a different structure (triangles).

mixed effects modelling for investigating the structure of a system (Pinheiro and Bates 2000) and various time-series analyses such as autoregressive components that deal with time effects (Brockwell and Davis 2002). Models which combine these aspects to integrate space, time and structure are sparse and only began to emerge recently in ecology (Banerjee et al. 2004, Wikle and Hooten 2010). Two key hurdles are that of estimating the uncertainties in the parameters of interest, and the computational time required to fit such models, especially for large data sets. The former can be overcome via Bayesian models through which the uncertainty is estimated as part of the fitting process, by treating all unknown quantities on a probabilistic basis. Solutions to the latter are an active topic of statistical research. Recently, Banerjee et al. (2008) have developed the predictive process (PP), a dimension-reduction technique for Kriging on large spatial data sets. The PP lowers the computational demands required to fit spatial models by reducing the dimension of the autocovariance matrix and thus enabling a richer set of models to be evaluated. The PP has not been widely used in ecology (though see Finley et al. 2009a, Latimer et al. 2009), especially for space-time data (though see Finley et al. 2012 which used the PP in a dynamic space-time model).

In this paper we apply the PP to one of the most frequent settings in ecology: multilevel spatio-temporal data. We extend the standard PP approach to consider more detailed hierarchical behaviour and include a separable temporal predictive process. We further correct for bias in the standard PP model by including the error correction terms proposed by Finley et al. (2009b). The main prediction here is that not considering the structure from which the data arises can result in misleading spatial patterns and consequently alter the way we interpret and act upon the results. We present our methodology by testing this hypothesis in a fisheries discards dataset. We investigate, among other questions, whether the hotspots of fisheries discards change with the inclusion of temporal effects or the various levels of the organizational structure of the data and explore which level or process is most responsible for this change. This can reveal for example if discards occur mainly due to the geographic location of the fish prone to be discarded or due to gear type, skipper behaviour or environmental conditions.

After describing the concept and importance of fisheries discards, we briefly describe the proposed methodology. One of the aims of this paper is to make this methodology more accessible to non-statisticians. As such, we clearly present the details of this approach by developing the models for our case study step by step and make the model code available to the readers. Subsequently, we explore and discuss the results of our case study and provide examples of other studies that can benefit from this approach.

Case study: fisheries discards

Discards are the portion of the catch from commercial fisheries returned to sea, and are estimated to be approximately 40% of the total catch in European waters (Kelleher 2005). This brings enormous conservation and ecological concerns as the survival rate of discarded organisms is extremely low (Kelleher 2005). Therefore, in order to sustainably manage the marine ecosystem it is crucial to mitigate discards. Understanding the pattern of discards is central to understanding the underlying causes and drivers of the discarding practice (Dunn et al. 2010), which is essential for developing mitigation tools. Indeed, spatial management of discarding has been recognized as one of the most successful discards reduction strategies (Roberts 1999). Discards data make an excellent case study, ecologically because very little is known about the spatial distribution of discards, and methodologically because when data are gathered it results in a specific organisational structure, common to many other ecological datasets.

We wish to determine if discarding events are spatially autocorrelated whilst taking account of their hierarchical structure. The structure of the relationship we propose is shown in Fig. 1. In Fig. 1a, each node corresponds to a discarding event located in space. Those closest to each other might share common features making them more similar than otherwise we would expect. In addition to the spatial information, discards have an inherently multilevel structure in which hauls are sampled from multiple trips that are performed by the same or different vessels (Fig. 1b). Therefore, features common to individual fishing trips representing, e.g. environmental conditions or mesh size, provide direct linkages among hauls, and features common to the vessel such as vessel size or gear type, may link discards rates among trips. Discarding events are also temporally explicit so that rates vary seasonally or from year to year (Viana et al. 2011). These temporal effects may be linked to spatial processes such as recruitment grounds and might provide linkages among trips represented by environmental conditions. In the latter, the influence of the sampling unit of trip could therefore act as a proxy for temporal variation. In any case, the overall observed pattern of discards is a product of the overlap between all these correlations processes (Fig. 1c).

There are several questions we might ask of such data in order to understand how each of these processes shape the spatial patterns of discards and ultimately how we should manage them: 1) where do the hotspots of discards occur and do they change after correcting for the organisational structure and time? 2) How do fisher-level factors, such as

behavior or vessel characteristics, influence spatial variation in discards, and how does this compare with environmental/ecological factors? 3) What is the spatial distance over which similar underlying causes of discards occur? All such questions involve disentangling the underlying processes.

Methods

Discards data

The Irish Marine Institute collects discards data on a regular basis through an onboard observers program. As discards are sampled from commercial fishing vessels, the available dataset has an inherent multilevel structure in which 430 hauls were sampled from 59 trips that were performed by 23 different vessels. With the purpose of illustrating the proposed methodology, here we model the natural log of discards per unit effort (or time spent hauling (dpue, kg h^{-1})) collected quarterly from 2003 to 2009 in the Irish Sea métier defined as 'Nephrops directed fishery using otter board trawls' (Nephrops métier, Fig. 2). Métier is a subdivision of a fishery defined with regards to the Data Collection Framework (EC Regulation 199/2008) which is based on landings percentage composition, area of operation and gear type. Thus, we investigate discards of the gear otter board trawl only. In addition to dpue, several other variables such as geographic coordinates, time, date, vessel name among many others are collected for each haul sampled.

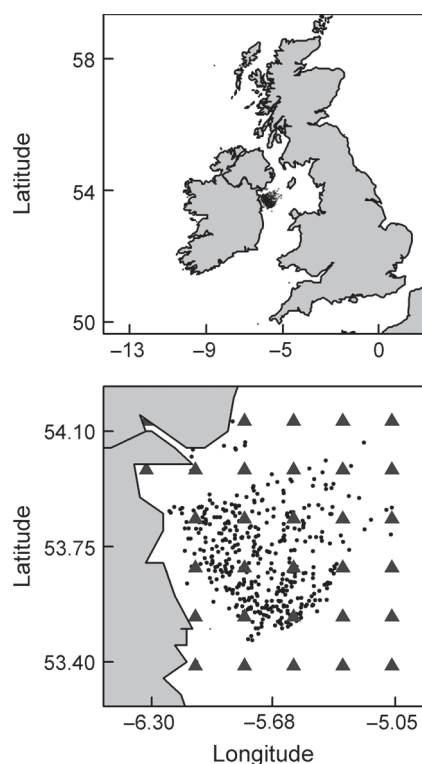


Figure 2. Irish Sea *Nephrops* directed fishery using otter board trawls métier (square in left panel) with the geographic locations of sampled hauls (points) from 2003 to 2008. The right panel shows an expanded version of the métier in which the dashed lines represent the 50 m bathymetry level and full line 100 m bathymetry level.

Hierarchical predictive process

The hierarchical PP was proposed by Banerjee et al. (2008), with minor corrections proposed by Finley et al. (2009b). This method provides a powerful tool for understanding the scale over which the underlying spatial environment exerts common pressures and shapes the overall pattern. It allows the use of spatially explicit models for very large spatial data sets, while simultaneously modelling multiple ecological processes that describe the spatial relationships among data points (Latimer et al. 2009).

The PP presented here is built under a Bayesian framework and is based on the kriging assumption that correlation among data points (spatial or temporal) decreases with distance. Most commonly, exponential decay functions are used to weight such correlations (for more detail see Diggle and Ribeiro 2002, Banerjee et al. 2004). The models we apply use this covariance structure as part of a Gaussian process applied to our response variable. This structure allows for fast and flexible model fitting, clear interpretation and simple interpolation techniques for prediction. The PP projects the original process observed at certain locations (usually a large number) onto a lower-dimensional subspace defined by the user. It is an approximation of a standard Gaussian process that depends on a far smaller set of locations, known as knots, on which the original process is trained. Figure 2 (bottom panel) shows the locations of the observations (black dots) with a grid of chosen knots overlaid (grey triangles). The PP will estimate the dpue in each of these knots rather than for each observed site. The dimension reduction enables far quicker calculation of the inverse autocorrelation matrix, required for evaluating the Gaussian process density. The larger original process is obtained via the usual Kriging stochastic interpolation. The number of knots should provide a good coverage of the process domain in order minimize the approximation error but it should also be as low as possible to keep the computational complexity low.

The main advantage of the methodology presented here is that we can combine this low demand method with other techniques that describe the remaining processes involved. Multilevel modeling has been increasingly applied in social (Snijders and Bosker 1999), medical (Langford et al. 1999) and more recently ecological fields (Wikle 2003, Wilson et al. 2010). In a multilevel structure, each level is fundamentally separated by different characteristics of the data or process. In the discards dataset this structure arises from the sampling and fishing activity i.e. hauls within trips and vessels. We also investigate the influence of temporal processes in the discarding practice of the Irish Sea by taking two different approaches: first we include time (per quarter year) as a quadratic fixed factor and second we use time under a spatio-temporal separable PP framework.

We define our response variable $Z_{jkq}(s, t)$ to be the natural log of dpue for location s , time t , haul j , trip k and vessel q . We write our model as:

$$Z_{jkq}(s, t) = \beta_0 + \beta_1 t + \beta_2 t^2 + w_1(s) + w_2(t) + u_k + v_q + \epsilon_{jkq} \quad (1)$$

where β_0 is an overall mean parameter, and β_1 and β_2 are the coefficients of linear and quadratic time respectively.

The PP is introduced via $w_1(s)$, a mean zero spatial PP, and $w_2(t)$, a mean zero temporal PP. The multilevel structure is defined via u_k , a random effect term for trip k so that $u_k \sim N(0, \sigma_u^2)$, v_q is a random effect term for vessel q so that $v_q \sim N(0, \sigma_v^2)$ and ϵ_{jkq} represents i.i.d model error so that $\epsilon_{jkq} \sim N(0, \sigma^2)$.

For the predictive processes, we define a set of k_s knots in space and k_t knots in time, here we use $k_s = k_t = 36$. The knots are chosen to cover the full range of locations and times and lie on a regular grid, and are set far fewer in number than the size of the data set n . We write the locations of the data as s_1, s_2, \dots, s_n and the times t_1, t_2, \dots, t_n . Similarly, we write the knots in space as $s_1^*, s_2^*, \dots, s_{k_s}^*$ and in time as $t_1^*, t_2^*, \dots, t_{k_t}^*$. We then define a Gaussian process with a exponential covariance on the spatial knots, written as $\tilde{w}_1(s^*) \sim GP(0, \sigma_s^2 \rho_s(s_i^*, s_j^*; \phi_s))$ where ρ_s is an autocorrelation function with $\rho_s(s_i^*, s_j^*; \phi_s) = \exp[-|d_{s_i^*, s_j^*}|/\phi_s]$, where $|d_{s_i^*, s_j^*}|$ is the distance between locations s_i^* and s_j^* (here the great circle distance), and a Gaussian process on the temporal knots, such that $\tilde{w}_2(t^*) \sim GP(0, \sigma_t^2 \rho_t(t_i^*, t_j^*; \phi_t))$ with $\rho_t(t_i^*, t_j^*; \phi_t) = \exp[-d_{t_i^*, t_j^*}^2/\phi_t]$. We follow Finley et al. (2009a, b) in correcting for bias in the predictive process by finally setting:

$$w_1(s) \sim N \left(\sum_{i=1}^{k_s} \sum_{j=1}^{k_s} \rho_s(s, s_i^*; \phi_s) c_{s,ij}^* \tilde{w}_1(s_j^*), \sigma_s^2 \left[1 - \sum_{i=1}^{k_s} \sum_{j=1}^{k_s} \rho_s(s, s_i^*; \phi_s) c_{s,ij}^* \rho_s(s, s_j^*; \phi_s) \right] \right) \quad (2)$$

$$w_2(t) \sim N \left(\sum_{i=1}^{k_t} \sum_{j=1}^{k_t} \rho_t(t, t_i^*; \phi_t) c_{t,ij}^* \tilde{w}_2(t_j^*), \sigma_t^2 \left[1 - \sum_{i=1}^{k_t} \sum_{j=1}^{k_t} \rho_t(t, t_i^*; \phi_t) c_{t,ij}^* \rho_t(t, t_j^*; \phi_t) \right] \right) \quad (3)$$

where $c_{s,ij}^*$ and $c_{t,ij}^*$ are the ij th elements of the precision matrix of the Gaussian processes $\tilde{w}_1(s^*)$ and $\tilde{w}_2(t^*)$ respectively. Note that this precision matrix is defined only on the k_s or k_t knots and is thus much faster to invert.

The multi-level trip and vessel effects allow us to disentangle the uncertainty in the discards created by repeated sampling from individual trips or vessels. For example, we are able to estimate and remove the variability of individual vessels, some of which may have particularly high or low discard rates. We can thus create more precise predictions. The key advantage of the spatial model is that precision is borrowed from neighbouring data points through the autocovariance function. The function can be thought of as containing a variance parameter σ_s^2 , which controls the variance of the spatial process, and a correlation function $\exp[-|d_{s_i^*, s_j^*}|/\phi_s]$ which determines the degree of correlation between data points according to their location. Here, ϕ_s controls the rate of decay as the distance d increases between locations s and s' . When the distance between sites is large the correlation decreases to 0. Oppositely, when the distance between sites is small, the correlation approaches 1. Note that we use an exponential covariance in space, but a Gaussian covariance in time as this is believed to be a much smoother process (Banerjee et al. 2004).

To test our hypothesis and investigate the different processes associated with discarding, we consider various spatial

and spatio-temporal models based on restricting Eq. 1. These restrictions correspond to Bayesian prior distributions on the various parameters which may be weak (e.g. giving a parameter a vague normal distribution) or strong (fixing a parameter to 0). All models investigated are described in Table 1. The model described in Eq. 1 is model I.

We fit such models via the Bayesian framework which allows for full parametric uncertainty and simple interpretation of results. We do not discuss the merits of Bayesian inference here but refer readers to Gelman et al. (2003) for further information. A Bayesian model will produce posterior probability distributions of the parameters given a likelihood (or data generating model) and a set of prior probability distributions for the parameters. In our case, the likelihood of the data is a multivariate normal distribution with mean and variance matrix determined by the multi-level spatial structure. We have relatively little prior information about the parameters for this particular problem, but we follow the approach of Finley et al. (2009a, b) to give inverse-Gamma priors to the residual variance $\sigma^2 \sim Ga(0.1, 0.1)$ and the spatial and temporal random effect variances $1/\sigma_s^2 \sim Ga(2, 1)$ and $1/\sigma_t^2 \sim Ga(2, 1)$. Our results are relatively robust to changes in this specification. The

autocorrelation parameters are often found to be hard to identify so we suggest slightly more informative priors $\phi_s \sim Ga(1, 0.1)$ and $\phi_t \sim Ga(7.5, 0.5)$. These values have been picked in order to constrain the autocorrelation to feasible values given the range of locations and times specific to our data. The former arises from the belief that a reasonable upper value for the correlation is to fall to 0.5 at a distance of around 100 km (the maximum distance between two points in our data). The latter prior corresponds to correlation falling to 0.5 at a maximum of 24 quarters (the maximum time-span between two points in our data). The 0.5 correlation value was chosen as the threshold between correlated and non-correlated. Thus, our prior rather generously allows at most a correlation of 0.5 between the end-members of our spatio-temporal data. Further details can be found in the Supplementary material Appendix 1.

We use the R (R Development Core Team) and JAGS (Plummer 2003) software to fit our models. The models are fitted by repeated Gibbs sampling of parameter values that match the likelihood and prior distributions. More detail can be found in Spiegelhalter et al. (1996). The algorithm requires starting values to begin, and thus can take time to converge to the posterior distribution. We run our model for 10^5 iterations, discarding 10^4 to achieve convergence. We use the Brooks, Gelman, Rubin diagnostic (Gelman and Rubin 1992) to check the status of convergence. More technical

Table 1. Description of all models (A to O) fitted with their prior distributions. We use the shorthand $\sim PP$ to indicate a parameter following a predictive process, and $\sim N$ to indicate a parameter following a normal distribution. β_1 and β_2 correspond to the coefficients of the quadratic effect on time, $w_1(s)$ and $w_2(t)$ correspond to the predictive process of space and time respectively, and u_k and v_q correspond to the random effects of trip and vessel, respectively. The hyper-parameters involved in these distributions are suppressed for clarity but can all be seen in Eq. 1.

| Model | β_1 | β_2 | $w_1(s)$ | $w_2(t)$ | u_k | v_q |
|-------|-----------|-----------|-----------|-----------|----------|----------|
| A | 0 | 0 | $\sim PP$ | 0 | 0 | 0 |
| B | 0 | 0 | $\sim PP$ | 0 | $\sim N$ | 0 |
| C | 0 | 0 | $\sim PP$ | 0 | $\sim N$ | $\sim N$ |
| D | $\sim N$ | $\sim N$ | $\sim PP$ | 0 | 0 | 0 |
| E | $\sim N$ | $\sim N$ | $\sim PP$ | 0 | $\sim N$ | 0 |
| F | $\sim N$ | $\sim N$ | $\sim PP$ | 0 | $\sim N$ | $\sim N$ |
| G | $\sim N$ | $\sim N$ | $\sim PP$ | $\sim PP$ | 0 | 0 |
| H | $\sim N$ | $\sim N$ | $\sim PP$ | $\sim PP$ | $\sim N$ | 0 |
| I | $\sim N$ | $\sim N$ | $\sim PP$ | $\sim PP$ | $\sim N$ | $\sim N$ |
| J | 0 | 0 | 0 | $\sim PP$ | 0 | 0 |
| K | 0 | 0 | 0 | $\sim PP$ | $\sim N$ | 0 |
| L | 0 | 0 | 0 | $\sim PP$ | $\sim N$ | $\sim N$ |
| M | $\sim N$ | $\sim N$ | 0 | $\sim PP$ | 0 | 0 |
| N | $\sim N$ | $\sim N$ | 0 | $\sim PP$ | $\sim N$ | 0 |
| O | $\sim N$ | $\sim N$ | 0 | $\sim PP$ | $\sim N$ | $\sim N$ |

details on the models and a JAGS code example can be found in Supplementary material Appendix 1.

We compare the different models by inspecting the posterior distributions of the variance parameters, noting their changing values between the different models. We further use the deviance information criterion (Spiegelhalter et al. 2002) which enables comparisons between models via the penalised expected deviance. The deviance is a measure of overall fit, whilst the penalty measures the complexity of the model by identifying the effective number of parameters in the model (Spiegelhalter et al. 1998). Owing to the large number of parameters used in the models, we use the 'popt' penalty from the r2jags package (Plummer 2003) to estimate the DIC which penalizes more severely the more complex models. A parsimonious model is expected to fit well and ideally will have relatively few parameters, thus a lower DIC can indicate a better model.

Once the model has been fitted, it is relatively simple to create predicted surfaces from a grid of new locations. The technical details of such predictions arise immediately from standard theory on the multivariate normal distribution. We can similarly sample uncertainties in the predicted surfaces from the posterior distributions of the parameters (see Supplementary material Appendix 2 for the method used to make the prediction surfaces).

Results

Based on the model selection criteria set out above, the most parsimonious model for describing the log of dpue in the Irish Sea *Nephrops* métier is model I. All models converged satisfactorily but the lowest DIC (Table 2) was obtained from model I, corresponding to a fully separable spatio-temporal model with all random effects included. The estimated values obtained from model I were found to

Table 2. Deviance information criterion (DIC) obtained from models A to O. The most parsimonious model, that with the lowest DIC, is highlighted bold.

| Model | DIC | Model | DIC |
|-------|-------|----------|--------------|
| A | 1018 | I | 917.6 |
| B | 920.0 | J | 991.1 |
| C | 921.8 | K | 923.8 |
| D | 1015 | L | 922.5 |
| E | 922.0 | M | 991.8 |
| F | 926.1 | N | 925.0 |
| G | 966.5 | O | 922.9 |
| H | 921.2 | | |

accurately describe the observed values as seen from the spatial distribution of the model variance, the normally distributed residuals and goodness of fit plots provided in Supplementary material Appendix 3, Fig. A1–A4. Our results were also robust to knot intensity and configuration. The full multilevel structure, i.e. hauls within trips within vessels, the fixed effect of time and the predictive effects of time and space are therefore necessary to understand and characterize the pattern associated with the log of dpue in the Irish Sea. Model B is the closest competitor for best model in terms of DIC (Table 2). However, the residuals were not as satisfactory as those from model I and since the best model in terms of DIC is still the most parameterized, even after using the 'popt' penalty which penalizes harsher more complex models, is a good indication of the goodness of the model chosen. Furthermore, model B does not include a temporal component but our results clearly support the presence of a temporal effect in discarding.

The log of dpue spatial pattern is shaped by the various processes underlying discarding as seen by the changing location, shape and intensity of the spatial residuals hot- (high values) and coldspots (lower values) in the predicted spatial surfaces (Fig. 3). The spatial surfaces in Fig. 3 correspond to the spatial random effects pattern that is not explained by any of the other parameters in the models, i.e. they show the hot- and coldspots of the fish. On the other hand, the spatial surfaces in Fig. 4 correspond to the mean discards distribution, i.e. Fig. 4A (obtained from model A) shows the discarding hotspots independently of the factors that are driving it, and Fig. 4I (the mean discards distribution of the best model, model I) shows the discards hot- and coldspots after removing the effect of vessel, trip and time. In our case study, the patterns observed for the spatial residuals distribution and the mean spatial distribution of discards are similar. If a residual hotspot remains after removing the effects of vessel, trip and time, this would mean that the fish hotspots match the discarding hotspots. As such, closing that area to fishing would be an appropriate mitigation measure. However, if after removing those effects there is no residual hotspot, this would mean that the fish hotspots do not match the discarding hotspots. As such, closing the area where the fish hotspots are, or the area where the discarding hotspots are, would only reallocate the drivers of discarding to a different area creating a discarding hotspot elsewhere rather than reducing discards. It is therefore of crucial to understand the residual spatial distribution.

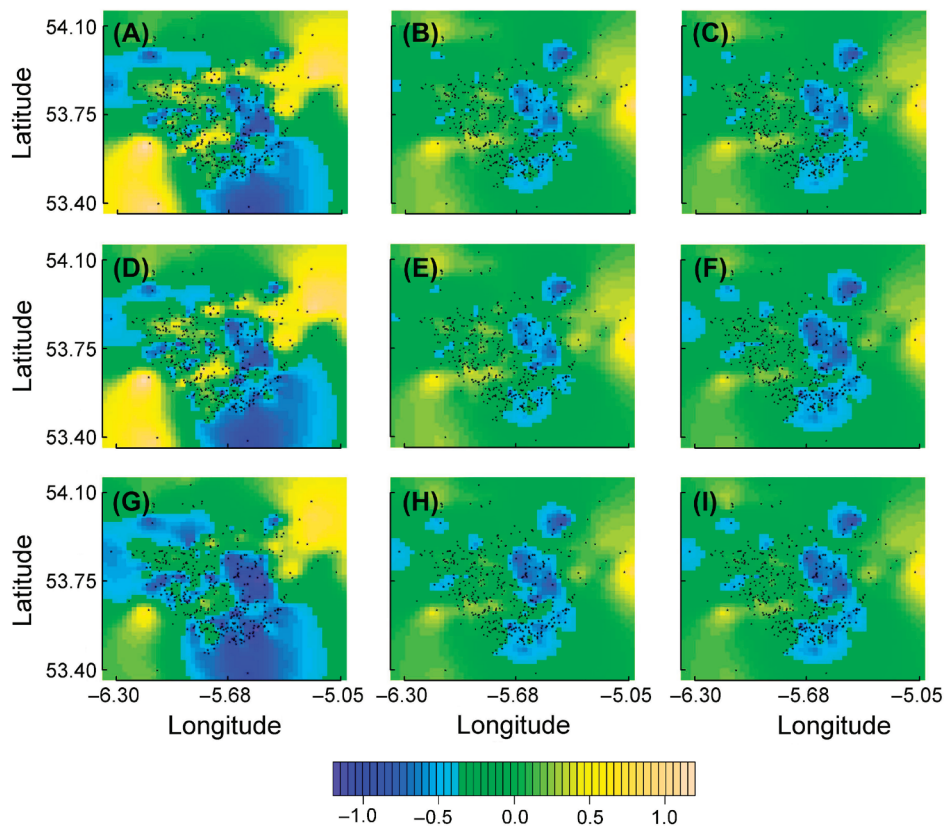


Figure 3. Spatial residuals prediction surface of model A: spatial PP without random effect; model B: spatial PP with trip as random effect; model C: spatial PP with trip and vessel as random effects; model D: spatial PP without random effects but with fixed effect of time; model E: spatial PP with trip as random effect and fixed effect of time; model F: spatial PP with trip and vessel as random effects and fixed effect of time; model G: spatial-temporal PP without random effects but with fixed effect of time; model H: spatial-temporal PP with trip as random effect and fixed effect of time; model I: spatial-temporal PP with trip and vessel as random effects and fixed effect of time.

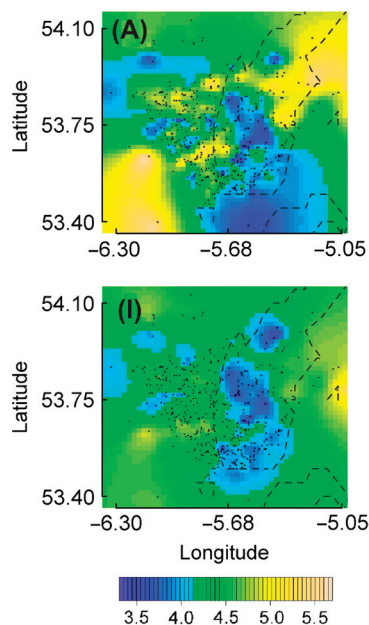


Figure 4. Overall spatial distribution (hotspots) of the log of discards per unit effort (dpue, kg h^{-1}) in the Irish Sea *Nephrops* metier. The dotted line represents the 75 m bathymetry line.

The multilevel structure of the data, in particular trip level, is the process that affects the spatial pattern the most. The inclusion of trip level (Fig. 3B, E, H) leads to the contraction of both the residuals hot and coldspots and also to the northward migration of the coldspot region. The inclusion of the second level of the sampling structure, i.e. vessel (Fig. 3C, F, I), is essential to characterize the discarding spatial process but seems to have little effect on the shape and pattern of the residual spatial surfaces. Similarly, inclusion of time as a linear predictor (Fig. 3D–I) has little effect on the spatial distribution, however time seems to slightly expand the residual coldspots (compared to model B and C without a linear predictor), this is particularly visible in the small isolated coldspots in the left hand side of Fig. 3C compared to Fig. 3F. Finally, the autocorrelation effect of time, described by the separable temporal PP (Fig. 3G–I) adds little or no changes to the residual spatial surface of log dpue in the Irish Sea. However, by investigating the effect of the temporal PP separately (Fig. 5) from the spatial distribution, we can understand its role in the discarding process and the reason it is necessary to describe the process.

The predictive effect of time reveals the change in discarding intensity over time. Independently of where discarding

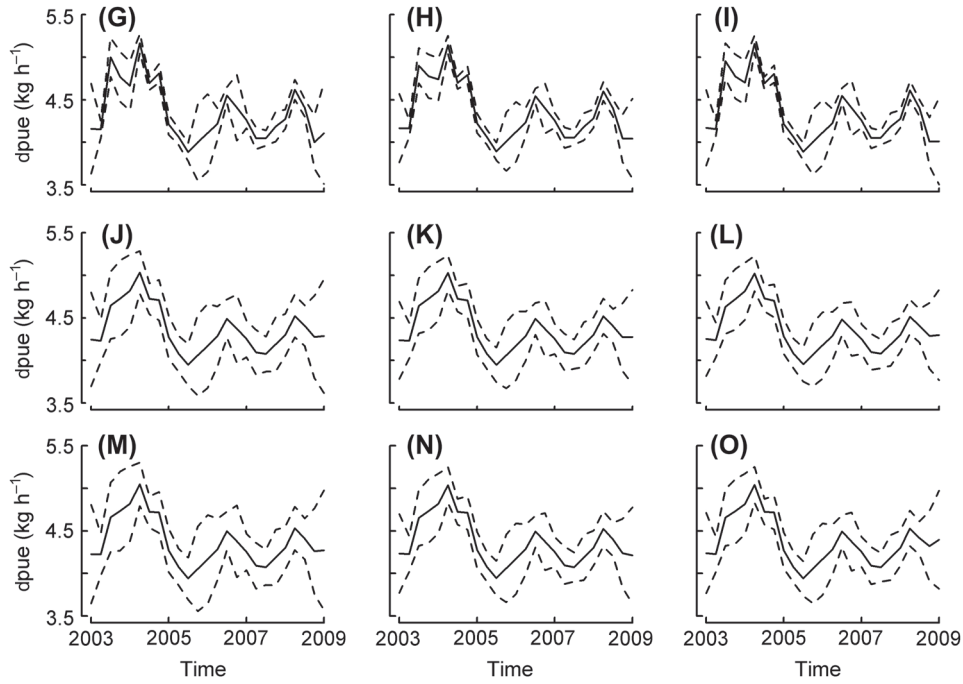


Figure 5. Temporal prediction of model G: spatial-temporal PP without random effects but with fixed effect of time; model H: spatial-temporal PP with trip as random effect and fixed effect of time; model I: spatial-temporal PP with trip and vessel as random effects and fixed effect of time; model J: temporal PP without random effects; model K: temporal PP with trip as random effect; model L: temporal PP with trip and vessel as random effects; model M: temporal PP without random effects but with fixed effect of time; model N: temporal PP with trip as random effect and fixed effect of time; model O: temporal PP with trip and vessel as random effects and fixed effect of time.

occurs, the log of dpue in the Irish Sea *Nephrops* métier has a strong cyclical behaviour of approximately two years as seen from Fig. 5. This indicates that following a year of low discarding there is a year of high discarding. Therefore, the temporal PP did not affect the overall mean of dpue (hence we could not detect any changes in the spatial distribution surfaces). Although small changes in the temporal process can be seen among models, the inclusion of one level or the full multilevel structure and the removal of the spatial process do not alter the observed temporal cycle.

To understand how the different processes influence each other, we quantify the uncertainty, estimated here as variance, of each parameter associated with each model

(see Fig. 6 and Supplementary material Appendix 4, Table A1 for the parameter values). When present, the spatial process (Fig. 6 models A–I, grey circles) and trip level (Fig. 6, black triangles) variances were consistently the highest (Fig. 6, grey circles) while vessel level (Fig. 6, grey triangles) have little or no variance. Most remarkably is that the inclusion of the temporal process not only explains part of the variance that was previously associated with trip level but also with the spatial process as seen from the decrease in these variances, from models without temporal process to models with the temporal process. The spatial variance in model I is lower than the variance associated with the temporal process (Fig. 6, black circles) indicating

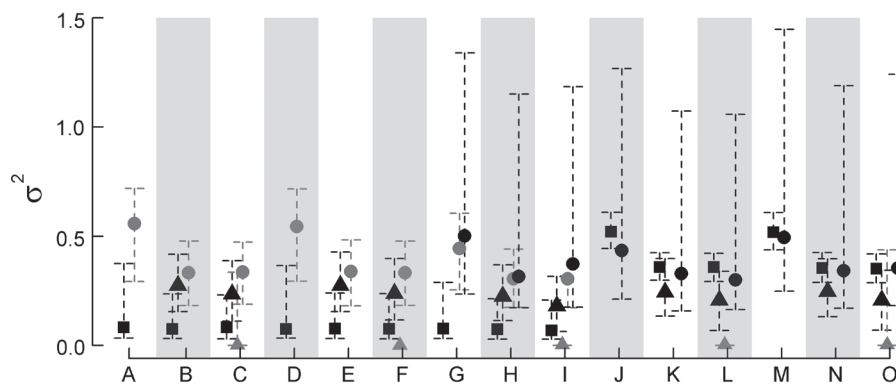


Figure 6. Variances associated with each process or level of model A to O. Black squares represent the residual variance; triangles represent the multilevel structure variance where black triangles represent trip level and grey triangles the vessel level variance; circles represent the predictive process variances where black circles represent the temporal term and grey circles the spatial term variance.

that most of the variance present in the log of dpue is due to a temporal process, more so than a spatial or organizational process. The residual variance (Fig. 6, black squares), i.e. the variance left to explain in the model, also drops with the inclusion of the temporal process but only when the spatial process is present. When the spatial process is removed the overall variance of the model increases considerably revealing the importance of the spatial process in the discarding process of the Irish Sea.

The exponential decay function that forms the kriging method gives information on the extent of spatial and temporal autocorrelation through the rate parameter ϕ (Supplementary material Appendix 4, Fig. A5). The spatial autocorrelation rate of decay (ϕ_s , Supplementary material Appendix 4, Fig. A5 upper panel) was highest in model I indicating a higher autocorrelation among discards events over long distances compared to other less complicated models. The temporal autocorrelation rate of decay of model I (Supplementary material Appendix 4, Fig. A5, bottom panel) was also among the highest of all models, similarly indicating a higher temporal autocorrelation between events separated by time than that of simpler models. However, all the spatial and temporal autocorrelation rates were very similar among the investigated models. The ϕ values, and all other estimated parameters for all models, can be found in the Supplementary material Appendix 4, Table A1.

Discussion

The proposed spatio-temporal model is a robust and versatile method for investigating the processes underlying the spatial patterns. Its main advantage lies in its ability to address both technical (e.g. autocorrelation and organisational structure) and practical problems (e.g. identifying hotspots, disentangling effects and describing basic spatial processes or quantification of uncertainties). While both multilevel modelling and geostatistics have been widely used in ecology and evolution, the PP method presented here has received little attention despite combining these two techniques into a versatile methodology with a low computational demand.

Spatial fisheries ecology has a direct applied relevance to resource management, but it also has a broad ecological significance (Botsford et al. 1997). Our results clearly support the prediction that hotspots are sensitive to the organisational structure of the data. If only the spatial autocorrelation process were considered in modelling the distribution of discards, drastic mitigation measures might be invoked. Two large hotspots, one in the southeast and another in northwestern part of the Irish Sea *Nephrops* métier were identified in model A which could drive a policy decision to close these areas for fishing. However, the inclusion of the trip and vessel level shows that the discarding hotspots are actually much smaller revealing that discarding and fish distribution are quite homogeneous in space. As a consequence, our results show that spatial closures based on high discarding locations are not an appropriate mitigation measure for the Irish Sea discards. Indeed, closing any area to fishing, in particular areas where fishing effort is high, would reduce

discards but would also proportionally reduce fishing profit and yield or create a discarding hotspot elsewhere, making this mitigation strategy unsuitable.

Similarly, the non-inclusion of the various underlying processes of discarding could have lead to the false conclusion that it is more sustainable to fish in the large area identified in model A as a discarding coldspot. However, part of the coldspot identified in model A is not actually a coldspot and therefore discarding in that area is higher than expected. Lower discarding seem to occur at water depths below the 75 m bathymetry line (Fig. 4) suggesting that the distribution of fish prone to be discarded is also shaped by bathymetry. Since discards are comprised mainly of fish below the minimum landings size, this concurs with Dickey-Collas et al. (1997) who found that small fish in the area are found in shallower waters around the 50 m bathymetry line.

Within the Irish Sea *Nephrops* métier, the temporal process seem to be key in the discarding process. Rather than mitigating discarding through closed areas, our results suggest that temporal closures might be a more appropriate management plan for the *Nephrops* fishery in the Irish Sea. Viana et al. (2011) suggested that in the Irish Sea higher discards of whiting, cod and haddock occurred during late spring and our results here show that discards have cycles of two years. Therefore, a mitigation strategy for this fishery that deserves further investigation is temporal closures during late spring occurring every other year. These 2-yr cycles could be driven by recruitment or spawning-stock biomass pulses in the species' population dynamics as shown by ICES (2009) for some species.

The spatial homogeneity of discarding can be partially explained by the relatively small autocorrelation between discarding events; after 10 km the spatial correlation of discards is already < 0.5 . This indicates that other processes are driving discards. Multilevel modelling is a way of breaking down complexity and disentangling the effects attributable to a number of different factors (Postma et al. 2012). In our case study, trip level was found to explain most of the variability previously associated with the haul level or even the spatial process. This indicates that factors associated with trip such as season, environmental conditions or mesh size, have a bigger influence in the discarding process of the Irish Sea *Nephrops* métier than the spatial process or the factors associated with vessel such as vessel characteristics and crew (Tamsett and Janacek 1999). This supports our suggestion that mitigation measures should target parameters such as season rather than vessel characteristics (e.g. vessel length and power). The organisational structure considers the links between the events which results in events from the same trip, during which for example high discarding occurred, being recognized as a product of that trip (or factors associated with) rather than a mere high occurrence of fish prone to be discarded. As a consequence the model takes this into account and removes this effect (e.g. resultant from skippers experience) from the process. The multilevel component of the model can therefore also be used to improve the sampling strategy and the results suggest that more trips within the same vessel should be sampled as there is a high variance associated with trip and very little with vessel.

Other ecological applications

There has been great interest in disentangling genetic and environmental effects on phenotype (Wolf et al. 1998, Postma et al. 2012). Our modelling framework can be applied to studies of this kind. For example, if we consider the genealogical tree of a population as the organisational structure of the data (imagine Fig. 1b represents the genealogical tree), and the distribution of phenotypes in space is known, we can separate the effects of spatial autocorrelation on phenotype (shared environmental effects) from effects due to genetic similarity.

Dispersal, invasion and conservation studies, such as the distribution of a species within a system of connected lakes (Granado-Lorencio et al. 2005) or obtained through stratified sampling (Latimer et al. 2009), can also benefit from this approach. In the first example, the organisational structure represents the connectivity of the lakes and we can tease apart the purely spatial and temporal variables, such as the geological characteristics of environment surrounding the lakes (closer lakes, even if not connected directly, might be more similar than those far apart), and those that follow the water flow e.g. eutrophication.

One more example that extends the range of topics amenable to analysis with our method involves species interactions. Looking at the ranges of two species, we might deduce that one is directly in competition with the other (Begon et al. 1996). However, including the organisational structure, in this case the food web in which they are embedded, may reveal that this effect occurs through indirect/apparent competition, e.g. an increase in species A drives the increase of predator abundance which in turn causes the decrease of species B, this means that if the food web was not considered in the analysis, the decrease in species B could have been attributed to for example niche competition with species A rather to increased predation.

Conclusions

While disentangling the underlying effects, valuable information was retrieved from the models: e.g. clearer and more precise prediction of hotspots, characterization of the spatial and temporal processes, insight on how to improve the sampling strategy and the recognition of the importance of factors associated with trip on discarding in the Irish Sea *Nephrops* métier. We confirmed our prediction that the organisational structure in which biological units are embedded alters the observed spatial pattern and concluded that in the fishery in question temporal processes are more important than spatial processes. This has profound consequences for management and conservation plans (as seen from the discards case study) as it influences our understanding of the system itself. The autocorrelation process and organisational structure are two of the most basic and common processes in ecological studies. However, many other processes can also influence the spatial patterns. The method proposed here serves as a base for more complex models as it allows for the inclusion of further processes, including single variables (e.g. temperature or depth) or functional forms (e.g. temporal dynamics; Langford et al. 1999).

However, this method also has its disadvantages. Owing to the relatively simplicity of the model, it is not possible to capture richer space-time dependence structures, namely how the hotspots move around with time. In the presence of a more comprehensive dataset, the next step in this methodology could then be to create a non-separable spatio-temporal model in which the spatial pattern is allowed to change location with time. Furthermore, if there is a high level of spatial detail that needs to be captured, the full Gaussian process might be preferable to the PP. Nonetheless, if one is interested in the links between emergent patterns and the underlying processes that drive them, the advantages of this simple and powerful framework for spatial ecology and conservation studies are evident.

Acknowledgements – This work (Grant-Aid Agreement no. PhD/FS/08/001) was carried out under the Sea Change strategy with the support of the Marine Inst. and the Marine Research Sub-programme of the National Development Plan 2007–2013 (Ireland).

References

- Banerjee, S. et al. 2004. Hierarchical modeling and analysis for spatial data. – Chapman and Hall/CRC.
- Banerjee, S. et al. 2008. Gaussian predictive process models for large spatial data sets. – *J. R. Stat. Soc. B* 70: 825–848.
- Begon, M. et al. 1996. Ecology: individuals, populations and communities, 3rd ed. – Blackwell.
- Botsford, L. W. et al. 1997. The management of fisheries and marine ecosystems. – *Science* 277: 509–515.
- Brockwell, P. J. and Davis, R. A. 2002. Introduction to time series and forecasting, 2nd ed. – Springer.
- Cressie, N. A. C. 1993. Statistics for spatial data. – Wiley.
- Dickey-Collas, M. et al. 1997. Does the western Irish Sea gyre influence the distribution of pelagic juvenile fish? – *J. Fish Biol.* 51: 206–229.
- Diggle, P. J. and Ribeiro, P. J. 2002. Bayesian inference in Gaussian model-based geostatistics. – *Geogr. Environ. Model.* 6: 129–146.
- Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – *Ecography* 30: 609–628.
- Dunn, D. C. et al. 2010. Spatio-temporal management of fisheries to reduce by-catch and increase fishing selectivity. – *Fish Fish.* 12: 110–119.
- Finley, A. O. et al. 2009a. Hierarchical spatial models for predicting tree species assemblages across large domains. – *Ann. Appl. Stat.* 3: 1052–1079.
- Finley, A. O. et al. 2009b. Improving the performance of predictive process modeling for large datasets. – *Comput. Stat. Data Anal.* 53: 2873–2884.
- Finley, A. O. et al. 2012. Bayesian dynamic modeling for large space-time datasets using Gaussian predictive processes. – *J. Geogr. Syst.* 14: 29–47.
- Gelfand, A. E. et al. 2005. Spatial process modelling for univariate and multivariate dynamic spatial data. – *Environmetrics* 16: 465–479.
- Gelman, A. and Rubin, D. 1992. Inference from iterative simulation using multiple sequences. – *Stat. Sci.* 7: 457–511.
- Gelman, A. et al. 2003. Bayesian data analysis, 2nd ed. – Chapman and Hall/CRC.
- Granado-Lorencio, C. et al. 2005. Abundance distribution relationships in fish assembly of the Amazonian floodplain lakes. – *Ecography* 28: 515–520.

- Gustafson, E. J. 1998. Quantifying landscape spatial pattern: what is the state of the art? – *Ecosystems* 1: 143–156.
- ICES 2009. Report of the ICES advisory committee 2009. – ICES Advice.
- Kelleher, K. 2005. Discards in the world's marine fisheries. An update. – FAO Fisheries Technical Paper 131.
- Langford, I. H. et al. 1999. Multilevel modelling of the geographical distributions of diseases. – *J. R. Stat. Soc. C* 48: 253–268.
- Latimer, A. M. et al. 2009. Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. – *Ecol. Lett.* 12: 144–154.
- Legendre, P. et al. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. – *Ecography* 25: 601–615.
- Levin, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. – *Ecology* 73: 1943–1967.
- MacArthur, R. H. 1984. *Geographical ecology: patterns in the distribution of species.* – Princeton Univ. Press.
- Pinheiro, J. C. and Bates, D. M. 2000. *Mixed effects models in S and S-Plus.* – Springer.
- Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. – In: Hornik, F. L. K. and Zeileis, A. (eds), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing.* Technische Univ. Wien, Vienna, Austria.
- Postma, E. et al. 2012. Disentangling the effect of genes, the environment and chance on sex ratio variation in a wild bird population. – *Proc. R. Soc. B* in press.
- Roberts, C. 1999. ACP-EU fisheries research report number 5 – marine protected areas as strategic tools. – RIVP, Netherlands.
- Snijders, T. A. B. and Bosker, R. J. 1999. *Multilevel analysis: an introduction to basic and advanced multilevel modeling.* – Sage Publishers.
- Spiegelhalter, D. et al. 1996. BUGS: Bayesian inference using Gibbs sampling, version 0.5. – MRC Biostatistics Unit, Cambridge.
- Spiegelhalter, D. J. et al. 1998. Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models. – Research Report 98-009, Division of Biostatistics, Univ. of Minnesota.
- Spiegelhalter, D. J. et al. 2002. Bayesian measures of model complexity and fit. – *J. R. Stat. Soc. B* 64: 583–639.
- Tamsett, D. and Janacek, G. 1999. Sampling trips for measuring discards in commercial fishing based on multilevel modelling of measurements in the North Sea from NE England. – *Fish. Res.* 42: 103–115.
- Urban, D. L. 2006. *Landscape ecology.* – Wiley.
- Viana, M. et al. 2011. Fishery discards in the Irish Sea exhibit temporal oscillations and trends reflecting underlying processes at an annual scale. – *ICES J. Mar. Sci.* 68: 221–227.
- Wikle, C. K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. – *Ecology* 84: 1382–1394.
- Wikle, C. K. and Hooten, M. B. 2010. A general science-based framework for spatio-temporal dynamical models. – *Test* 19: 417–451.
- Wilson, T. L. et al. 2010. Hierarchical spatial models for predicting pygmy rabbit distribution and relative abundance. – *J. Appl. Ecol.* 47: 401–409.
- Wolf, J. B. et al. 1998. Evolutionary consequences of indirect genetic effects. – *Trends Ecol. Evol.* 13: 64–69.

Supplementary material (Appendix E7853 at <www.oikosoffice.lu.se/appendix>). Appendix 1–4.