Ecography

E7853

Viana, M., Jackson, A. L., Graham, N. and Parnell, A. C. 2012. Disentangling spatio-temporal processes in a hierarchical system: a case study in fisheries discards. – Ecography 35: xxx–xxx.

Supplementary material

Supplementary material

Appendix 1: Technical details

Notation:

$Z_{jkq}(s,t)$	log of dpue for site s, time t, haul j, trip k, vessel q
s	Spatial location
s^*	Spatial knots
t	Time
t^*	Temporal knots
u_k	Random effect for trip
v_q	Random effect for vessel
$w_1(s)$	Predictive process in space
$w_2(t)$	Predictive process in time
β	Linear regression parameters
$\epsilon_{jkq}(s.t)$	Residual/model error
ρ_s	Spatial exponential autocorrelation function
ρ_t	Temporal exponential autocorrelation function
σ^2	Residual variance
σ_s^2	Spatial random effect variance
σ_t^2	Temporal random effect variance
σ_u^2	Trip random effect variance
σ_v^2	Vessel random effect variance
ϕ_s	Spatial autocorrelation scale parameter
ϕ_t	Temporal autocorrelation scale parameter
$C(s_a, s_b; \phi_s)$	$= \sigma_s^2 \rho(s_a, s_b; \phi_s)$. Spatial autocovariance between sites a and b
$H(t_a, t_b; \phi_t)$	$= \sigma_t^2 \rho(t_a, t_b; \phi_s)$. Temporal autocovariance between times a and b

Model description:

We have a model defined as:

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

with the following prior distributions:

- $u_k \sim N(0, \sigma_u^2)$
- $v_k \sim N(0, \sigma_v^2)$
- $\epsilon_{jkq}(s,t) \sim N(0, \sigma^2)$
- $w_1(s) \sim PP(0, \sigma_s^2 \rho(s_i, s_j; \phi_s))$
- $w_2(t) \sim PP(0, \sigma_t^2 \rho(t_i, t_j; \phi_t))$

Here PP is a predictive process, representing a low-rank version of a Gaussian process. We use the Finlay et al. (2009) corrected version so that:

$$w_1(s) = C(s, s^*; \phi_s) C(s^*, s^*; \phi_s)^{-1} \tilde{w}_1(s^*) + \tilde{\epsilon}_s$$

with $\tilde{w}_1(s^*)$ a standard Gaussian Process with exponential covariance defined on knots s^* and $\tilde{\epsilon}_s \sim N(0, diag(C(s, s; \phi_s) - C(s, s^*; \phi_s) C(s^*, s^*; \phi_s)^{-1} C(s^*, s; \phi_s)))$ where $diag$ takes only the diagonal elements and sets the remainder of the matrix to 0. This correction term is required to adjust the variance of the predictive process and thus avoid over-smoothing. Note that its effect is to adjust that variance matrix of w so that the diagonals are all σ_s^2 but the off-diagonals remain unchanged. Analogously for the temporal model we have:

$$w_2(t) = H(t, t^*; \phi_t) H(t^*, t^*; \phi_t)^{-1} \tilde{w}_2(t^*) + \tilde{\epsilon}_t$$

with $\tilde{w}_2(t^*)$ a standard Gaussian Process with exponential covariance defined on knots t^* and $\tilde{\epsilon}_t \sim N(0, diag(H(t, t; \phi_t) - H(t, t^*; \phi_t) H(t^*, t^*; \phi_t)^{-1} H(t^*, t; \phi_t)))$. The key advantage of the predictive process is that we only ever require the inverse of an $s^* \times s^*$ (or $t^* \times t^*$) matrix vastly reducing the computational burden.

JAGS code:

```
model{
    for(i in 1:N){
        Z[i] ~ dnorm(mu[i], prec[i])
        mu[i] = beta0 + beta1*t[i] + beta2*pow(t[i],2) + u[k[i]] + v[q[i]] + w.tilde[i] + w.t.tilde[i]
        prec[i] = 1/var.all[i]
        var.all[i] = tausq + sigmasq - correction[i] + sigmasq.t - correction.t[i]
        }
# Spatial Predictive process
    w.tilde.star ~ dmnorm(mu.w.star, C.star.inv)
    C.star.inv = inverse(C.star)
    for(i in 1:N.star) {
        mu.w.star[i] = 0
        C.star[i,i] = sigmasq
        for(j in 1:(i-1)) {
            C.star[i,j] = sigmasq*exp(-(d.s.star.star[i,j]/phi))
            C.star[j,i] = C.star[i,j]
      } }
# Interpolate spatial PP back on to original sites
    for(i in 1:N) {
        for(j in 1:N.star) {
            C.s.star[i,j] = sigmasq*exp(-(d.s.star[i,j]/phi))
      } }
    w.tilde = C.s.star%*%C.star.inv%*%w.tilde.star
# Temporal Predictive process
    w.t.tilde.star ~ dmnorm(mu.w.t.star,H.star.inv)
    H.star.inv = inverse(H.star)
    for(i in 1:N.t.star) {
        mu.w.t.star[i] = 0
        H.star[i,i] = sigmasq.t
        for(j in 1:(i-1)) {
            H.star[i,j] = sigmasq.t*exp(-(d.t.star.star[i,j]/phi.t))
            H.star[j,i] = H.star[i,j]
      } }
# Interpolate temporal PP back on to original sites
    for(i in 1:N) {
        for(j in 1:N.t.star) {
            H.t.star[i,j] = sigmasq.t*exp(-(d.t.star[i,j]/phi.t))
      } }
    w.t.tilde = H.t.star%*%H.star.inv%*%w.t.tilde.star
# Variance correction
    for(i in 1:N) {
        correction[i] = t(C.s.star[i,])%*%C.star.inv%*%C.s.star[i,]
        correction.t[i] = t(H.t.star[i,])%*%H.star.inv%*%H.t.star[i,]
    }
```

```
# Prior distributions
tausq = 1/tausq.inv
tausq.inv ~ dgamma(0.1,0.1)
sigmasq = 1/sigmasq.inv
sigmasq.inv ~ dgamma(2,1)
phi ~ dgamma(1,0.1)
beta0 ~ dnorm(0,0.0001)
beta1 ~ dnorm(0,0.0001)
beta2 ~ dnorm(0,0.0001)
sigmasq.t = 1/sigmasq.t.inv
sigmasq.t.inv ~ dgamma(2,1)
phi.t ~ dunif(7.5,0.5)

# Prior distributions for random effects
for(i in 1:Nk){ u[i] ~ dnorm(0,tau.u) }
tau.u= 1/(sigma.u*sigma.u)
sigma.u ~ dunif(0,1000)

for(i in 1:Nq){ u[i] ~ dnorm(0,tau.v) }
tau.v= 1/(sigma.v*sigma.v)
sigma.v ~ dunif(0,1000)

} #end model
```

where, *d.s.star.star* and *d.s.star* are the geographic distances between spatial knots and between these and the observed locations respectively (in kilometers), and *d.t.star.star* and *d.t.star* are the distances between the temporal knots and between these and the observed time points respectively. To decrease running time, we calculated the distances in R and passed them on to JAGS as data.

Appendix 2: Prediction surface methods

The technical details for making a prediction arise immediately from standard theory on the multivariate normal distribution. We find that writing our own R code produces quicker results than incorporating the prediction surface in JAGS while running the models themselves. We consider that the prediction of any Z follows equation S1 in which the notation $*$ refers to the predicted parameters. Suppose now we wish to predict new observations Z_0 for given new locations s_0 and new times t_0. We make use of a standard result from the multivariate normal distribution, namely that if:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

then,

$$x_1|x_2 = a \sim N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right).$$

Whilst the PP is useful for estimating the parameters of the autocovariance function we do not use it for prediction as there is little computational burden in using the full Gaussian Process (GP) model. In terms of results, in this case study it makes very little difference whether we use the GP or PP for prediction hence we chose the method that is less demanding computationally. Thus in our scenario we have $x_1 = Z_0$, $x_2 = Z$ and $\mu_1 = \mu_2 = \beta_0$, thus:

$$\begin{aligned}
\Sigma_{11} = Var(Z_0) &= var(w_1(s_0)) + var(w_2(t_0)) + var(u_k) + var(v_q) + var(\epsilon) \\
&= C_{s_0} + H_{t_0} + \Sigma_u + \Sigma_v + \tau^2 I \\
\Sigma_{22} = Var(Z) &= var(w_1(s)) + var(w_2(t)) + var(u_k) + var(v_q) + var(\epsilon) \\
&= C_s + H_t + \Sigma_u + \Sigma_v + \tau^2 I \\
\Sigma_{21} = \Sigma_{12}^T = Cov(Z, Z_0) &= C(s, s_0; \phi_s) + H(t, t_0; \phi_t)
\end{aligned}$$

Predictions for the reduced models or for the spatial random effect w can be produced from the above by removing the appropriate terms.

Appendix 3: Model fit of most parsimonious model (model I)

A model can be more parsimonious than others and still be inappropriate to describe the investigated process. In this section of the supplementary information, we show the goodness of fit of our most parsimonious model (model) which was found to accurately describe the log of dpue the Irish Sea *Nephrops* métier. Fig. S1 shows the spatial distribution of the residual variance. If a model is appropriate we should find low variance where samples were taken and high variance as the prediction moves away from the observed location. Model I, follows this pattern (Fig. S1).

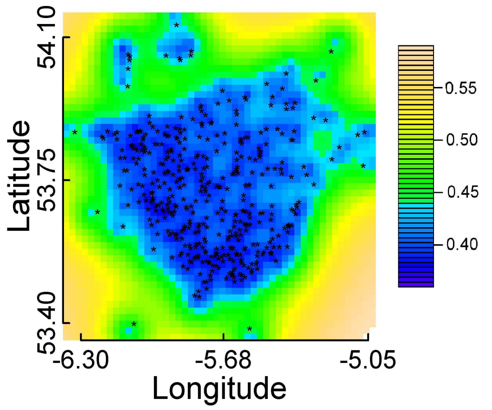

Figure A1: Spatial distribution of the residual variance for model I. The dots represent the locations of the samples.

Since the model should predict the observed values, the relationship between the observed and estimated values of a parsimonious model should result in a slope close to 1. Figure S2 shows the relationship between the observed and estimated log of dpue for model I. This confirms that the selected model, model I, is in fact a parsimonious model and suitably describes the log of dpue in the *Nephrops* métier of the Irish Sea as the correlation between the observed and estimates values is 0.7.

A good model is also expected to have independent normal distributed residuals. In a quantile-quantile plot (Q-Q plot) the linearity of the points suggests that the data are normally distributed. The distribution of residuals of model I (Figure S3), our chosen model, showed normally distributed residuals suggesting the model is parsimonious.

A model that suitably describes the data should also predict the relationship of the observed values with the various parameters involved in the model specification. Our

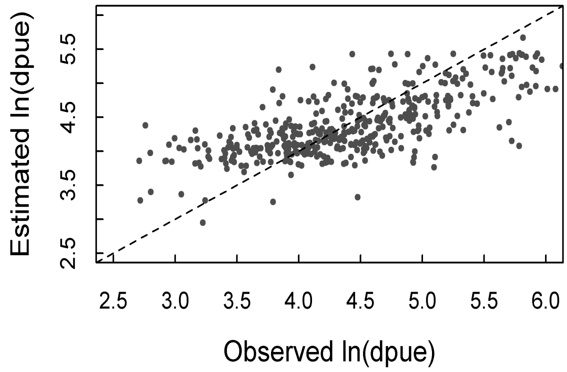

Figure A2: Log of the observed discards per unit effort (dpue, kg.h^{-1}) against estimated log of dpue for model I. The dotted line represents a perfect fit of intercept 0 and slope 1.

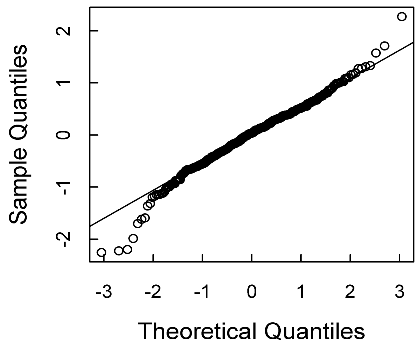

Figure A3: Q-Q plot of the residuals of model I, our chosen model.

chosen model, model I, also passed this test as seen from Figure S4 that shows that the estimated mean values (blue dots) over time match the relationship of the observed mean values with time.

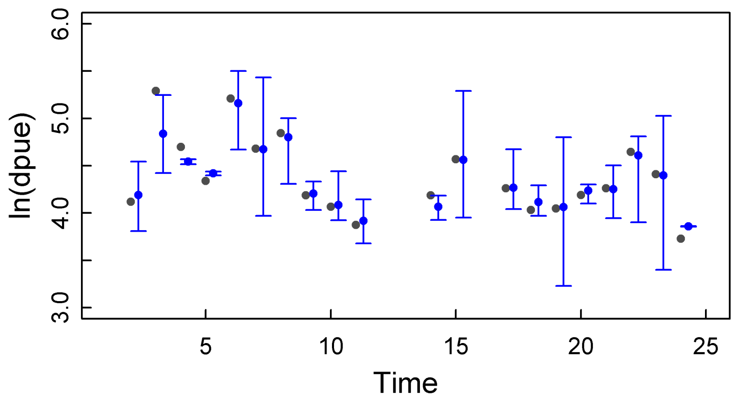

Figure A4: Observed (grey dots) and estimated mean values with respective 95% credible intervals (blue dots)against time. The estimated values are offset in the x-axis to make visualization clearer.

Appendix 4: Supplementary results

Table S1, here split into Table S1a and S1b, shows the coefficient modes and respective 95% credible intervals of all estimated parameters for all models, model A to model O.

Table A1a: Coefficient modes with 95% standard errors of Models A to O. Where, β_0 is the intercept, β are the coefficients of the quadratic term of time and ϕ and ϕ_2 are the rates of decay for the spatial and temporal components respectively.

	β_0	β_1	β_2	ϕ_s	ϕ_t
A	4.506		-	21.962	-
	(4.36, 4.77)			(13.74, 42.82)	
B	4.375	-	-	24.110	-
	(4.15, 4.67)			(9.59, 54.12)	
C	4.418	-	-	25.512	-
	(4.16, 4.73)			(11.07, 53.82)	
D	4.832	-0.030	0.001	22.192	-
	(4.47, 5.24)	(-0.08, 0.03)	(-0.00, 0.00)	(14.19, 44.46)	
E	4.599	-0.016	0.000	24.773	-
	(3.94, 5.25)	(-0.13, 0.10)	(-0.00, 0.00)	(10.13, 56.45)	
F	4.642	-0.022	0.001	24.926	-
	(3.94, 5.40)	(-0.15, 0.10)	(-0.00, 0.00)	(9.64, 54.66)	
G	5.549	0.019	0.000	23.998	13.732
	(2.59, 6.21)	(-0.24, 0.27)	(-0.01, 0.01)	(15.77, 50.98)	(6.82, 26.49)
H	4.391	0.017	-0.001	24.849	15.366
	(2.51, 6.02)	(-0.22, 0.28)	(-0.01, 0.01)	(10.16, 55.46)	(7.54, 28.00)
I	4.440	0.006	-0.000	26.222	14.736
	(2.56, 6.09)	(-0.24, 0.27)	(-0.01, 0.01)	(9.60, 56.32)	(7.48, 28.01)
J	4.432	-	-	-	12.736
	(3.25, 5.56)				(6.52, 26.64)
K	4.321	-	-	-	14.465
	(3.07, 5.37)				(7.52, 29.00)
L	4.428	-	-	-	14.534
	(3.48, 4.39)				(7.13, 28.44)
M	4.334	0.004	-0.000	-	12.917
	(2.34, 6.15)	(-0.26, 0.27)	(-0.01, 0.01)		(6.98, 26.19)
N	4.389	0.007	-0.001	-	14.279
	(2.68, 6.06)	(-0.24, 0.26)	(-0.01, 0.01)		(7.09, 27.70)
O	4.391	-0.006	0.001	-	14.130
	(2.49, 6.23)	(-0.25, 0.26)	(-0.01, 0.01)		(7.30, 28.15)

Table A1b: Variance mode with 95% standard errors of Models A to O. Where, σ^2 is the overall model variance and $\sigma^2_{spatial}$, $\sigma^2_{temporal}$, σ^2_{trip} and σ^2_{vessel} are the spatial, temporal, trip and vessel level variances, respectively.

	σ^2	$\sigma^2_{spatial}$	σ^2_{time}	σ^2_{trip}	σ^2_{vessel}
A	0.084	0.557	-	-	-
	(0.03, 0.36)	(0.29, 0.72)			
B	0.075	0.332	-	0.275	-
	(0.03, 0.24)	(0.18, 0.48)		(0.15, 0.42)	
C	0.083	0.335	-	0.235	0.002
	(0.03, 0.23)	(0.19, 0.47)		(0.11, 0.39)	(0.00, 0.34)
D	0.076	0.544	-	-	-
	(0.03, 0.37)	(0.29, 0.72)			
E	0.077	0.339	-	0.275	-
	(0.03, 0.24)	(0.18, 0.48)		(0.15, 0.43)	
F	0.076	0.333	-	0.238	0.000
	(0.03, 0.24)	(0.18, 0.48)		(0.12, 0.40)	(0.00, 0.03)
G	0.078	0.443	0.501	-	-
	(0.03, 0.29)	(0.25, 0.60)	(0.23, 1.34)		
H	0.074	0.305	0.316	0.225	-
	(0.03, 0.21)	(0.17, 0.44)	(0.17, 1.15)	(0.11, 0.37)	
I	0.070	0.305	0.374	0.182	0.003
	(0.03, 0.21)	(0.17, 0.44)	(0.17, 1.18)	(0.06, 0.32)	(0.00, 0.37)
J	0.520	-	0.434	-	-
	(0.44, 0.61)		(0.21, 1.27)		
K	0.359	-	0.330	0.243	-
	(0.30, 0.43)		(0.16, 1.07)	(0.13, 0.40)	
L	0.359	-	0.300	0.207	0.003
	(0.30, 0.42)		(0.16, 1.06)	(0.07, 0.34)	(0.00, 0.39)
M	0.518	-	0.493	-	-
	(0.44, 0.61)		(0.25, 1.45)		
N	0.355	-	0.342	0.248	-
	(0.29, 0.43)		(0.17, 1.19)	(0.13, 0.40)	
O	0.352	-	0.353	0.209	0.003
	(0.29, 0.42)		(0.18, 1.24)	(0.07, 0.34)	(0.00, 0.44)

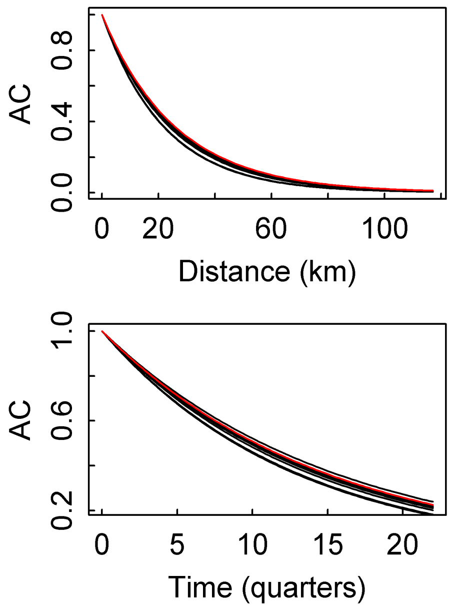

Figure A5: Posterior mean of the spatial (left panel) and temporal (right panel) autocorrelation rate (AC) of decay estimated from all models. Red line highlights model I ($\phi_s = 26.2$ and $\phi_t = 14.7$).