

# Statistical Reports

*Ecology*, 100(1), 2019, e02403  
 © 2018 by the Ecological Society of America

## Black swans in space: modeling spatiotemporal processes with extremes

SEAN C. ANDERSON<sup>1,2,4</sup> AND ERIC J. WARD<sup>3</sup>

<sup>1</sup>School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, Washington 98195 USA

<sup>2</sup>Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia V6T 6N7 Canada

<sup>3</sup>Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanographic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, Washington 98112 USA

**Citation:** Anderson, S. C., and E. J. Ward. 2019. Black swans in space: modeling spatiotemporal processes with extremes. *Ecology* 100(1):e02403. 10.1002/ecy.2403

**Abstract.** In ecological systems, extremes can happen in time, such as population crashes, or in space, such as rapid range contractions. However, current methods for joint inference about temporal and spatial dynamics (e.g., spatiotemporal modeling with Gaussian random fields) may perform poorly when underlying processes include extreme events. Here we introduce a model that allows for extremes to occur simultaneously in time and space. Our model is a Bayesian predictive-process GLMM (generalized linear mixed-effects model) that uses a multivariate-*t* distribution to describe spatial random effects. The approach is easily implemented with our flexible R package *glmmfields*. First, using simulated data, we demonstrate the ability to recapture spatiotemporal extremes, and explore the consequences of fitting models that ignore such extremes. Second, we predict tree mortality from mountain pine beetle (*Dendroctonus ponderosae*) outbreaks in the U.S. Pacific Northwest over the last 16 yr. We show that our approach provides more accurate and precise predictions compared to traditional spatiotemporal models when extremes are present. Our R package makes these models accessible to a wide range of ecologists and scientists in other disciplines interested in fitting spatiotemporal GLMMs, with and without extremes.

**Key words:** Bayesian statistics; ecological extremes; geostatistical models; heavy-tailed distributions; mountain pine beetle; multivariate-*t* distribution; random fields; spatial statistics; spatiotemporal models; Stan.

### INTRODUCTION

Applications of statistical models that allow for joint inference about spatial and temporal dynamics have advanced rapidly in ecology over the last several decades (e.g., Bascompte and Solé 1995, Latimer et al. 2009, Conn et al. 2015). Spatiotemporal models have also been widely used in other disciplines, including applications to weather, remote sensing, human disease dynamics, and crime (Cressie and Wikle 2011). When ecological data are spatially structured, explicitly accounting for spatial autocorrelation can improve model predictions and inference about parameters of interest (e.g., Shelton et al. 2014, Thorson et al. 2015, Ver Hoef et al. 2017).

Manuscript received 12 October 2017; revised 26 January 2018; accepted 29 March 2018; final version received 24 May 2018. Corresponding Editor: Derek Johnson.

<sup>4</sup>E-mail: sean.anderson@dfo-mpo.gc.ca

Including spatial components in statistical models involves extending models that most ecologists are familiar with, such as generalized linear models (GLMs) or generalized additive models (GAMs). Spatial relationships can be included as predictors in models of the mean (e.g., a two-dimensional GAM) or can be included in models of the covariance (e.g., kriging). Recent extensions of these spatial covariance models include modeling spatial deviations in GLMs as random effects (Gaussian random fields, GRFs). GRFs represent a two-dimensional version of Gaussian processes and define the expected value, variance, and covariance of random draws from a multivariate normal distribution. Examples of these methods in existing R packages include, but are not limited to, *spBayes* (Finley et al. 2007), *INLA* (Rue et al. 2009), and *spate* (Sigrist et al. 2015) (Appendix S1: Table S1).

A limitation of spatial models that use GRFs is that they may perform poorly when data include extreme or

“black-swan” events. Black-swan events refer to rare and seemingly improbable events that nonetheless happen, often with large consequences (Taleb 2007). Recent work has demonstrated the regular occurrence of such events in natural and ecological processes (e.g., Albeverio et al. 2006, Ward et al. 2007, Fey et al. 2015, Anderson et al. 2017). When models describing a spatial process also include an observation model, for example, anomalous observations may be reconciled by increasing the variance of the observation error (rather than attributing these to extremes in the process). Extremes in temporal processes have been modeled using a variety of methods in ecology, typically by including mixtures of normal and heavy-tailed distributions (e.g., Everitt 1996, Ward et al. 2007, Thorson et al. 2011). More recently, the Student *t* distribution has been proposed as a solution to modeling process variation with extremes, or black-swan events, in population dynamics (Anderson et al. 2017).

Several extensions of GRFs have been proposed to better capture extreme spatial events, including max-stable or extreme value theory (Davison and Gholamrezaee 2012, Davison et al. 2012), where quantities of interest include probabilities of exceeding some threshold (Davis and Mikosch 2008). Other extensions of spatiotemporal models to describe extremes include the use of multivariate-*t* (MVT) spatial random fields (Røislien and Omre 2007). Compared to GRFs, MVT random fields allow for increased flexibility in the representation of spatial patterns through time while converging on the multivariate normal (MVN) GRFs when appropriate (Røislien and Omre 2007).

In this paper, we introduce the use of robust spatial predictive models using the MVT distribution, and provide a user-friendly implementation in our R package *glmmfields*. Using simulation testing, we illustrate that the MVT model leads to better prediction (greater accuracy, more precision) when the spatial process includes heavy-tailed events. As an application to real-world data, we apply this model to data on mountain pine beetle (*Dendroctonus ponderosae*) outbreaks in the Pacific Northwest of the United States.

## METHODS

We seek to allow for large deviations in an ecological spatial pattern over time by extending spatiotemporal predictive process models to use a MVT distribution instead of a MVN distribution. Below we describe the form of the model as implemented in *glmmfields*, describe two simulation tests exploring model performance, and finally describe the application of our model to a data set of mountain pine beetle outbreaks.

### *Predictive process models*

Latimer et al. (2009) provide an overview of predictive process models for ecologists. For large data sets, estimating spatial random effects at many locations may be

computationally prohibitive. One solution is to estimate a spatial field as correlated random effects at a subset of locations or *m* “knots” (e.g., Latimer et al. 2009, Shelton et al. 2014), where *m* < *n*, the number of data points (Appendix S1: Fig. S1). The location of the *m* knots describing a random field can be chosen via a clustering algorithm, such as the partitioning around medoids algorithm (Reynolds et al. 2006). Instead of estimating an unconstrained *m* × *m* covariance matrix, a covariance function is specified a priori to model covariance as a function of distance. Our *glmmfields* package allows for the isotropic squared exponential (Gaussian), exponential, and Matern covariance functions; however, anisotropic functions could be included in the future. Given estimated random effects at the knot locations and the known distance matrix between the knots and observed data, the knot random effects can be projected to the locations of the observations (Røislien and Omre 2007, Latimer et al. 2009, Finley et al. 2009, Appendix S1: Fig. S1).

As an example, the squared exponential covariance function models the correlation between points *i* and *j* as  $H(\delta_{ij}) = \exp(-\delta_{ij}^2/2\theta_{GP})$ , where  $\delta_{ij}$  is the distance between points *i* and *j* and  $\theta_{GP}$  controls how steeply correlation declines with distance (GP, gaussian process). For a given set of  $\delta_{ij}$ , large values of  $\theta_{GP}$  correspond to smooth spatial patterns and small values correspondence to wiggly spatial patterns. The elements of the covariance matrix  $\Sigma$  at the *m* knot locations are then defined as  $\Sigma_{ij}^* = \sigma_{GP}^2 \exp(-\delta_{ij}^2/2\theta_{GP})$  with the spatial variance parameter  $\sigma_{GP}^2$  scaling the amplitude of the spatial deviations and the \* denoting a symbol referring to the knot locations as opposed to the sample locations. Following Latimer et al. (2009), we can calculate the covariance matrix  $\Sigma_{(W,W^*)}$  between the spatial random effects  $W$  at the sample locations and the realizations of the spatial process  $W^*$  at the knot locations. Given  $\Sigma^*$ , we can generate  $W^*$  by drawing from a multivariate distribution (MVN or MVT) with covariance  $\Sigma^*$  and projecting these to the data locations as  $\Sigma_{(W,W^*)}: W = \Sigma'_{(W,W^*)} \Sigma^{*-1} W^*$ .

### *MVT random fields*

Our model is essentially a GLMM with a spatiotemporal element described by a MVN or MVT random field. We define the mean of location *s* and time *t* as  $\mu_{s,t} \equiv \mathbb{E}(y_{s,t})$ . We then define  $g(\mu_{s,t}) = \mathbf{X}_{s,t}\beta + \gamma_{s,t}$ , where  $g$  is a link function,  $\mathbf{X}_{s,t}$  represents predictors, and  $\beta$  represents a vector of estimated coefficients. The symbol  $\gamma_{s,t}$  represents the spatiotemporal process, described below. The variance of the observation model depends on the chosen error distribution (e.g., Gaussian, Poisson, or gamma).

Modifying the MVN spatiotemporal model to the more flexible MVT distribution requires estimating the degrees of freedom parameter *v*. When *v* is small (say *v* < 10) the distribution has heavier tails than the MVN, meaning, extreme events are more likely (Fig. 1). For most purposes, the MVT and MVN are

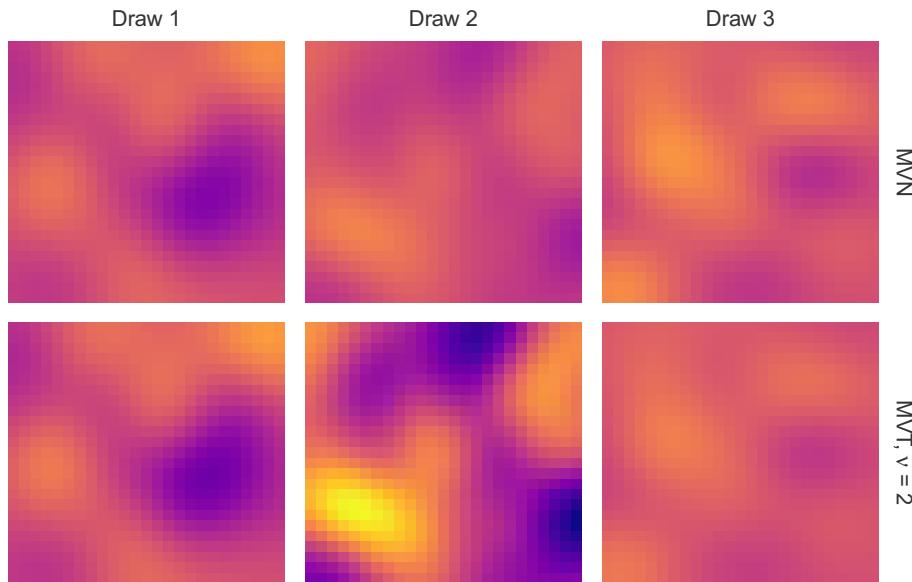


FIG. 1. An illustration of three draws from a multivariate normal (MVN) random field (top row) vs. three draws from a multivariate-t (MVT) random field with heavy tails (degrees of freedom,  $v$ , of 2; bottom row). Draws represent slices of time with independent spatial processes at each time slice. The scale parameters  $\sigma_{GP}$ , spatial decay parameters  $\theta_{GP}$ , random seeds, and colour scales are held constant between the two rows to illustrate the differences. Note the considerably more extreme values (darker purples and yellows) in the second draw for the MVT spatial process.

indistinguishable for moderate values of  $v$  (say  $v > 20$ ) similarly to the univariate  $t$ -distribution compared with the normal distribution (e.g., Anderson et al. 2017). If  $W_{s,t}$  defines the value of a random field at spatial location  $s$  and time  $t$ , then the spatiotemporal element  $\gamma_{s,t}$  can be made temporally constant (one field shared across time,  $\gamma_{s,t} = W_s$ ), independent at each time step ( $\gamma_{s,t} = W_{s,t}$ ), or autoregressive so that the spatial pattern at time  $t$  is dependent on the spatial pattern at time  $t - 1$  to a degree defined by  $\phi$ , ( $\gamma_{s,t} = \phi\gamma_{s,t-1} + W_{s,t}$ ).

Our glmmfields package fits these models in a Bayesian framework. We sample from the posterior distribution using the No-U-Turn Sampler, which is an extension of Hamiltonian Markov Chain Monte Carlo (MCMC), implemented in Stan (Stan Development Team 2016b, Carpenter 2017) and the R package rstan (Stan Development Team 2016a). This Bayesian approach has a number of advantages. First, it lets us fully quantify uncertainty around all parameter estimates and derived quantities (e.g., predictions, probabilities of exceeding extremes). Second, the Bayesian framework lets us place weakly informative priors on parameters to impose our existing knowledge of reasonable values and to aid computation. In the case of the degrees of freedom parameter,  $v$ , we bound the lower value to 2 for computational stability and use a gamma (shape = 2, rate = 0.1) prior, which has a mean of 20 and a median of about 17 (Juárez and Steel 2010).

#### Testing the recovery of spatial extremeness

We used simulation testing to evaluate how well heavy spatial tails could be recovered under various conditions.

Our simulations included 50 data points (15 knots) collected at the same locations annually for 5, 15, or 25 yr. The spatial process was independent by year ( $\gamma_{s,t} = W_{s,t}$ ) with  $\sigma_{GP}^2 = 1$  (the spatial variance) and  $\theta_{GP} = 1$  (the spatial correlation parameter). We set the MVT  $v$  parameter to 2.5 (heavy tails), 5 (moderately heavy tails), or 20 (effectively normal tails). We then corrupted the “true” spatial process with observation error, using a gamma distribution with a log link,  $y_{s,t} \sim \text{gamma}(\text{shape} = a, \text{rate} = a/\mathbb{E}(y_{s,t}))$ , where the shape parameter  $a$  can be reparameterized into the coefficient of variation (CV),  $a = 1/\text{CV}^2$ . We tested CVs of 0.1, 0.6, and 1.2. The underlying linear predictor  $\mathbf{X}_{s,t}\boldsymbol{\beta}$  was set to zero to focus on the spatial process. Therefore, the simulation of data  $y_{s,t}$  simplifies to  $\log(\mu_t) \sim \text{MVT}(v, 0, \Sigma_W)$ ,  $y_{s,t} \sim \text{gamma}(1/\text{CV}_{\text{obs}}^2, 1/\text{CV}_{\text{obs}}^2 \cdot \mu_{s,t})$ . We attempted to recover  $v$  by fitting a model that matched the process generating the simulated data. We provide a full description of the priors and joint probability distributions along with details of the MCMC sampling for this and all other models in Appendix S1.

#### Testing the advantage of allowing for extremes

To evaluate the consequence of assuming spatial processes are not present when they actually are, we generated simulated data sets from a model with spatial extremes (MVT) and compared the fit of models with and without extremes (MVT vs. MVN). Specifically, we simulated data from the model:  $\mu_t \sim \text{MVT}(v, 0, \Sigma_W)$ ,  $y_{s,t} \sim \text{normal}(\mu_{s,t}, \sigma_{\text{obs}})$ , with  $\sigma_{GP} = 0.3$  (the scale of the spatial deviations),  $\theta_{GP} = 1.2$  (the spatial correlation

parameter),  $\sigma_{\text{obs}} = 0.8$ , and 100 spatial data points. We set the MVT  $v$  parameter to 2 to represent very heavy tails. The locations of the data, and the knots (15), were held constant through time to enable faster computations.

To evaluate out-of-sample predictive accuracy, we withheld 10% of the data randomly (10 points per year) from the model fitting and then compared the MVT and MVN models. We compared the root mean squared error between the  $\log(\hat{\mu}_{s,t}^{\text{withheld}})$  posterior medians and the true  $\log(\mu_{s,t}^{\text{withheld}})$ , the width of the 95% credible intervals (CIs) on  $\hat{\mu}_{s,t}^{\text{withheld}}$ , and the difference in the leave-one-out information criteria (LOOIC), a Bayesian information criteria that approximates leave-one-out predictive performance (Vehtari et al. 2017, Appendix S1).

#### *Mountain pine beetles in the U.S. Pacific Northwest*

To illustrate real-world applications of spatial models with extremes, we fit MVT and MVN random field models to a data set representing mountain pine beetle outbreaks in the U.S. Pacific Northwest from 1994 to 2014 (USDA Forest Service 2017). We rasterized the map into a 500 by 500 grid and then aggregated this high-resolution grid into percent cover in a coarser grid reduced by a factor of 25. We modelled the proportion of finer grid cells affected by outbreaks per coarser grid cell. We excluded a small number of coarser cells without outbreaks since we chose to fit this example with a lognormal observation model for simplicity. An alternative, to include the zeros, would be a beta-binomial observation model. Since the proportion affected was far from 1, we can fit a model with a log link and lognormal observation distribution. The log of the mean proportion

affected at location  $s$  and time  $t$ ,  $\mu_{s,t}$ , is predicted by a year-specific random walk defined by  $\beta_t$  and the spatiotemporal process  $\gamma_{s,t}$ , with the spatial process itself modelled as autoregressive:

$$\log(\mu_{s,t}) = \beta_t + \gamma_{s,t}, \quad (1)$$

$$\beta_t \sim \text{normal}(\beta_{t-1}, \sigma_\beta), \quad (2)$$

$$\gamma_t^* \sim \text{MVT}(v, \phi \gamma_{t-1}^*, \Sigma_W^*). \quad (3)$$

We model the data,  $y_{s,t}$ , as generated by a lognormal observation model with scale parameter  $\sigma_{\text{obs}}$ :  $y_{s,t} \sim \text{lognormal}(\log(\mu_{s,t}), \sigma_{\text{obs}})$ .

We used 20 knots to represent the spatial process (increasing the number of knots did not substantially affect the results). We compared the above MVT model to a Gaussian random field model. To evaluate out-of-sample predictive accuracy we withheld 25 randomly selected data points per year from the model fitting, for a total of 400 withheld data points, or approximately 10% of the data. We then compared the log predictive density for the held-out data and the width of the 95% CIs (details, including priors in Appendix S1).

#### RESULTS

Under most scenarios we were able to recapture true values of  $v$  with reasonable accuracy and low bias (Fig. 2, Appendix S1; Fig. S2). The number of time steps had the largest effect on detecting low values of  $v$ . For example, the median absolute proportional error between  $v$  and  $\hat{v}$  (median of the posterior) was only 0.18 with 25 time steps, minimal observation error (CV = 0.1), and  $v = 2.5$  (Fig. 2b). However, the median

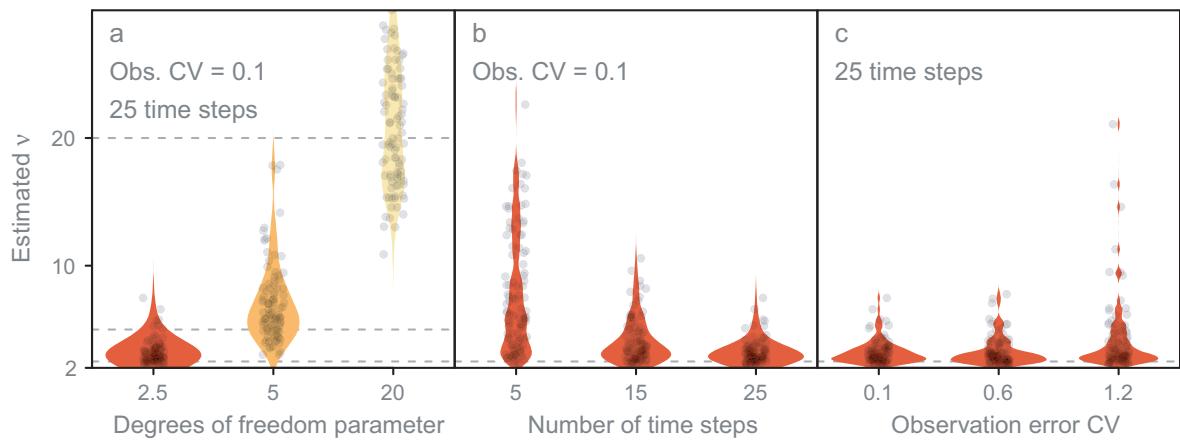


FIG. 2. Simulation testing the ability to recapture the degree of spatial heavy tailedness in an MVT random fields model. Shown are tests with (a) various true values of  $v$  (the MVT degrees of freedom parameter), (b) an increasing number of time steps in the data set (with low observation error), and (c) an increasing level of observation error. The full factorial results are shown in Appendix S1: Fig. S2. Individual dots show the median estimates from individual simulation runs. The color scale indicates the true degree of heavy tailedness from yellow (effectively normal) to red (very heavy tailed) and the violin-shaped polygons represent the density.

absolute proportional error increased by approximately two- and eightfold when the number of time steps was reduced to 15 and 5, respectively (Fig. 2b). Observation error did not substantially affect the estimation of  $v$  until relatively high levels of observation error (i.e., CV = 1.2, Fig. 2c).

When the true underlying data were generated with spatiotemporal extremes (MVT,  $v = 2$ ), fitting a model without extremes reduced out-of-sample predictive accuracy and precision (Fig. 3). The MVN model tended to overestimate  $\sigma_{GP}$  (which controls the magnitude of the random field deviations) to account for effectively fixing  $v = \infty$  (e.g., Fig. 3a). The out-of-sample RMSE (root mean squared error) was slightly higher for the misspecified MVN model compared to a correctly specified MVT model (median RMSE = 0.24 vs. 0.23; Fig. 3b). The out-of-sample MVN model CIs were a median of 11% larger than the MVT model CIs. The leave-one-out information criterion (LOOIC) correctly chose the MVT model in 96% of the simulations (Fig. 3c). Separately, if linear predictors were included, we found them to be largely unbiased under the misspecified MVN model: the effect of misspecification was on the random field estimation and therefore the predictions in space and time.

For the pine beetle case study, the spatial model with extremes generated more accurate and precise out-of-sample predictions (Fig. 4). The estimates of  $v$  indicated heavy tails in the spatiotemporal process (median  $v = 2.4$ , 95% CI = 2.0–3.9, Fig. 4c). Advantages of the MVT model over the MVN included greater log predictive density for held-out data (Fig. 4d) and a median of 35% narrower 95% CIs. Predictions from the MVT model demonstrate evolving hotspots of pine beetle infestation

in the U.S. Pacific Northwest, with particularly strong hotspots in 2009–2010 (Fig. 5).

## DISCUSSION

We have introduced a spatial process model that models random fields through time with a MVT rather than MVN distribution. Through simulation, we demonstrate the advantages of this new model: when spatiotemporal extremes exist in simulated data, our MVT model has superior predictive accuracy and precision. Using a case study on mountain pine beetle outbreaks in the Pacific Northwest, we show that our MVT model produces more accurate and precise out-of-sample predictions compared to a MVN model. Because the MVT model converges to the MVN model and only requires estimating one additional parameter, we recommend fitting the MVT model even if anomalous events are not thought to be present a priori.

Many ecological, environmental, and anthropogenic processes could generate spatiotemporal extremes. For example, unmodelled animal movement such as changes to the shoaling behaviour of fish with population density declines (Rose and Kulka 1999) could generate apparent extremes. Another possible cause could be spatial climate anomalies. For instance, a series of abnormally warm winters around 1990 led to a sudden northward shift in juvenile cod followed by low abundance in the southern North Sea (Rindorf and Lewy 2006). In southern Australia, several weeks of extreme warm water temperatures led to the sudden range contraction of a brown alga seaweed (Smale and Wernberg 2013). Heat waves can also increase the likelihood of wildfires or disease outbreaks (e.g., Harvell et al. 2002),

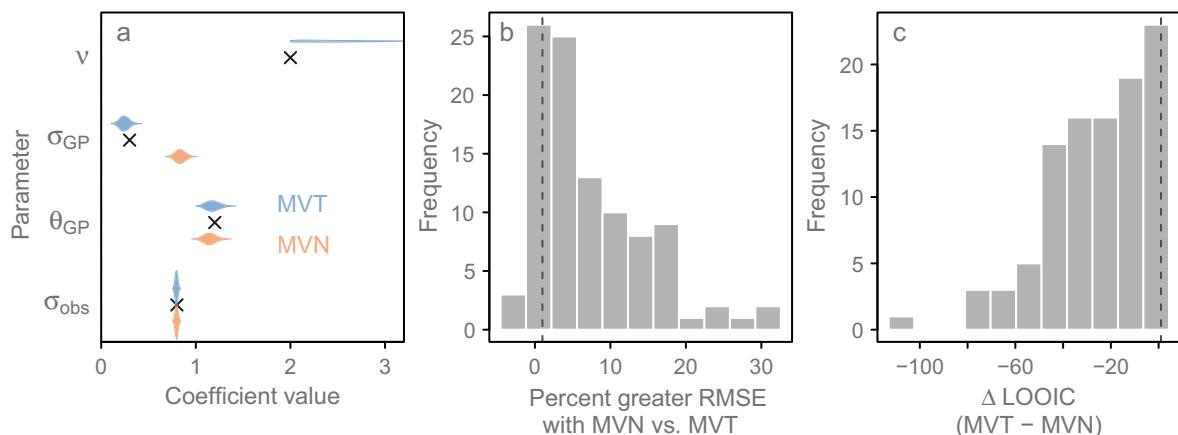


FIG. 3. The effect of fitting MVN and MVT models when the true underlying data are drawn from simulated MVT random fields. (a) Parameter posteriors for an example model fit of a mismatched MVN model (orange) and a correct MVT model (blue). Black crosses indicate true values. The parameter  $v$  is challenging to estimate but can generally distinguish between providing evidence for heavy tails (much lower than the prior) and not providing evidence for heavy tails (matching the prior; Anderson et al. 2017). The parameter  $\sigma_{GP}$  represents the spatial standard deviation,  $\theta_{GP}$  represents the spatial decay or correlation parameter, and  $\sigma_{obs}$  represents the observation error standard deviation. (b) Percent greater root mean squared error (RMSE) and (c)  $\Delta$  LOOIC (leave-one-out information criteria) for the mismatched MVN model compared to the correct MVT model (negative  $\Delta$  LOOIC values favor the MVT model).

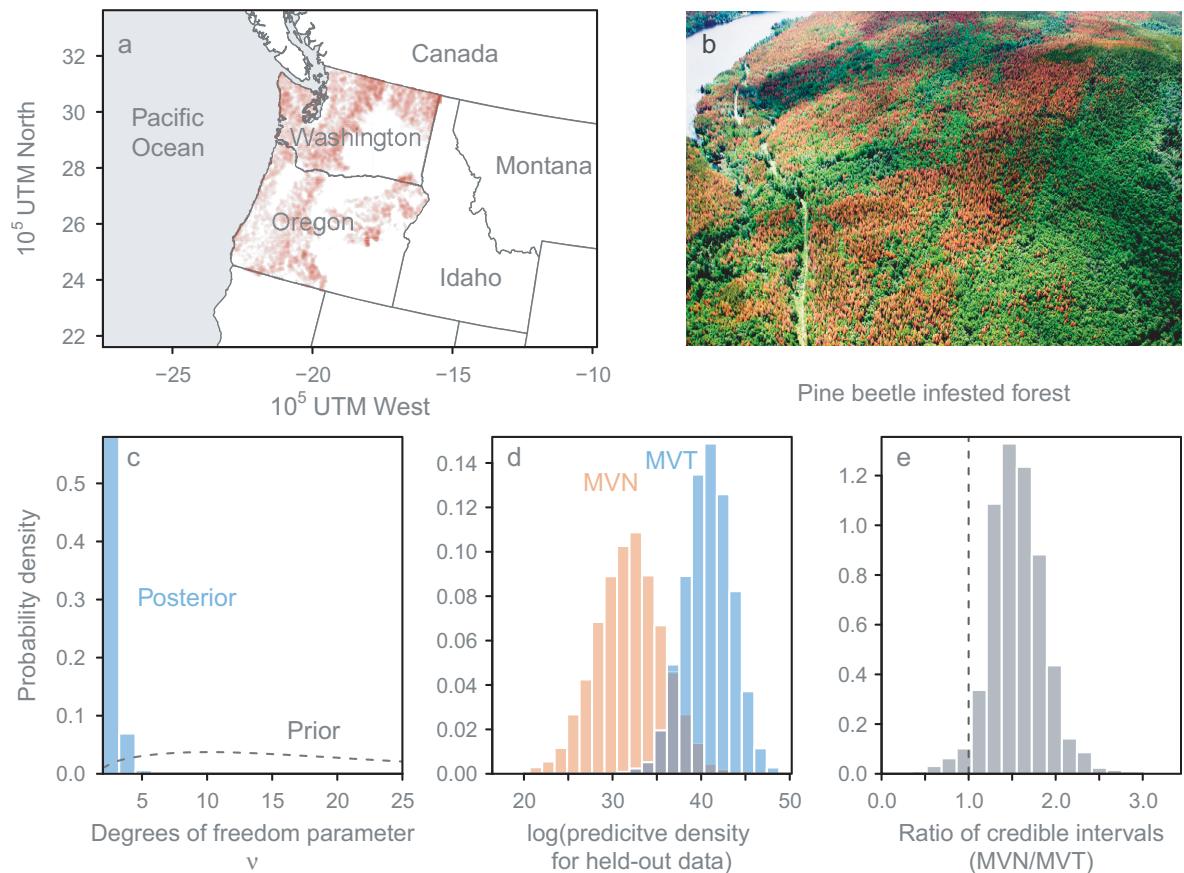


FIG. 4. Model fit characteristics of autoregressive random field models fit to mountain pine beetle data from the U.S. Pacific Northwest. (a) Map of the region with tree mortality observations for 1 yr shown with brown dots. (b) Photograph of a pine beetle infested forest in British Columbia, Canada. Photo by Dezene Huber. (c) Posterior and prior distributions of the degrees of freedom parameter,  $v$ . Low  $v$  values (approximately  $v < 10$ ) indicate evidence of heavy tails. (d) Log predictive density for held-out data for MVT (blue) and MVN (orange) models. (e) The ratio of 95% CI widths on the predicted percent beetle cover in time and space between the MVN and MVT models (i.e., on the predictions shown in Fig. 5).

which could generate spatial extremes through time. Additionally, human-caused disasters, such as marine oil spills could cause extremes in spatial patterns through time. In essence, the MVT random fields model allows for unexpected and unmodelled events to occur and have less influence on parameter estimates and predictions.

For our case study of pine beetle outbreaks, the MVT model provided better out-of-sample predictive ability and greater precision compared to the same model with an MVN random field. We did not link covariates, such as temperature, to the mean response, but it is possible that their inclusion could help explain the spatial anomalies. Importantly, our model is descriptive rather than dynamic (Cressie and Wikle 2011). A dynamic model would include mechanisms governing the spatial evolution of the pine beetle outbreaks (such as available tree host size and quality; Chubaty et al. 2009), and these types of mechanisms could generate the extremes we observe with our descriptive model. Alternatively,

incorporating spatiotemporal extremes into a dynamic model of pine beetle outbreaks might improve its predictive capacity.

Our associated R package, *glmmfields*, is designed to be familiar for anyone who has fit GLMs in R, and includes flexibility for many features beyond those described so far. For example, *glmmfields* can be used to fit spatial GLMMs for data without a temporal component (Appendix S2). The package can fit observation models beyond those included in this paper, such as Poisson, negative binomial, or binomial, and can include numeric or factor covariates. *glmmfields* includes familiar R utility functions (e.g., *predict()*) and plotting functions for model checking. Furthermore, the package includes a simulation function that can simulate data from any model the package can fit, and a series of unit tests that simulate and fit all major model configurations.

There are limitations to the approach we have described. First, spatial predictive models require the

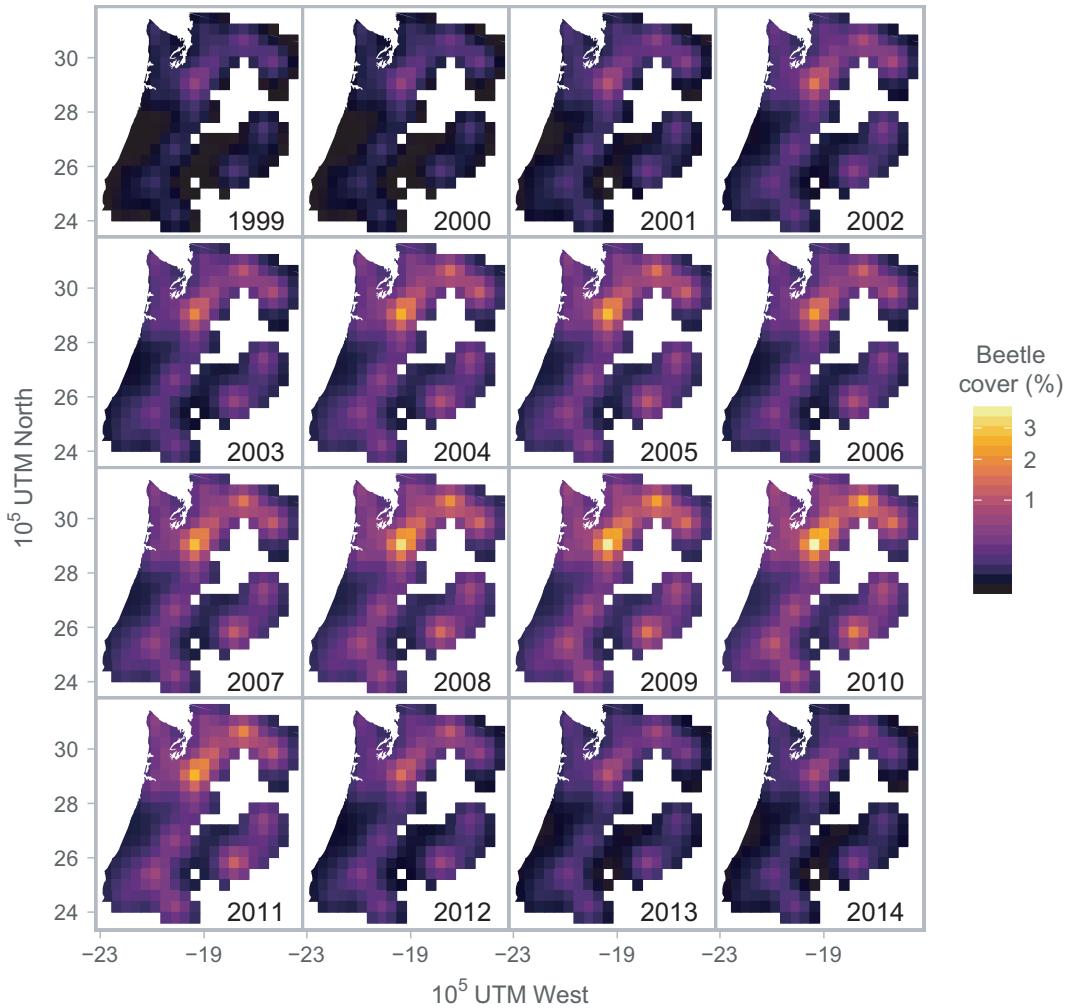


FIG. 5. Modeled severity of mountain pine beetle outbreaks in Washington and Oregon, USA from 1999 to 2014. Shown are medians of the modeled parameter  $\mu_{s,t}$  from Eq. 1: an autoregressive spatiotemporal MVT random fields model. The color scale is square-root distributed.

selection of knots; too few may not characterize the spatial field accurately, biasing parameter estimates. The spatial process may also be poorly described by an MVT or MVN random field and be better described by some other form (e.g., see Conn et al. 2015). While the MVT spatial model is more robust than the MVN model, better predictions and inference might be obtained by explicitly modeling the processes that generated the extremes. Also, the MVT random fields model requires a sufficient number of observations and sufficiently low sampling error to detect spatial extremes (though when extremes cannot be detected, the model converges to the MVN model).

Our MVT spatial model uses a predictive approach to achieve considerable efficiency over modeling a full covariance matrix describing all observed locations. However, an important topic for research is to include sparse matrix algorithms in spatiotemporal models of

extremes. A recent advance to spatial models with large datasets has been the stochastic partial differential equation (SPDE) approximation to GRFs proposed by Lindgren et al. (2011). These methods are accessible via the integrated nested Laplace approximation (INLA; Rue et al. 2009), which allows for approximate Bayesian sampling of the posterior without MCMC. Use of the SPDE-INLA approach has increased rapidly in ecology over the last 5 yr (e.g., Illian et al. 2013, Ono et al. 2016), and is significantly faster than other approaches, in part because of integration with Template Model Builder through software such as VAST (Thorson and Barnett 2017). Regardless, the MVT random field model introduced here is already efficient, accessible to a wide range of ecologists through the included R package, and allows us to improve predictions for ecological processes with extreme spatial anomalies through time.

## ACKNOWLEDGMENTS

We thank T. A. Branch and J. T. Thorson for helpful discussions on the modeling approach and B. J. Harvey for helpful advice on the mountain pine beetle dataset. We are grateful to the maintainers of the U.S. Forest Service Insect and Disease Survey Database. We thank J. T. Thorson and A. M. Edwards for helpful comments on an earlier version of this manuscript. Funding was provided by a David H. Smith Conservation Research Fellowship to S. C. Anderson.

## LITERATURE CITED

- Albeverio, S., V. Jentsch, and H. Kantz, editors. 2006. Extreme events in nature and society. Frontiers collection. Springer, Berlin, Germany.
- Anderson, S. C., T. A. Branch, A. B. Cooper, and N. K. Dulvy. 2017. Black-swan events in animal populations. *Proceedings of the National Academy of Sciences USA* 114:3252–3257.
- Bascompte, J., and R. V. Solé. 1995. Rethinking complexity: modeling spatiotemporal dynamics in ecology. *Trends in Ecology and Evolution* 10:361–366.
- Carpenter, B. 2017. Stan: a probabilistic programming language. *Journal of Statistical Software* 76:1–32.
- Chubaty, A. M., B. D. Roitberg, and C. Li. 2009. A dynamic host selection model for mountain pine beetle, *Dendroctonus ponderosae* Hopkins. *Ecological Modelling* 220:1241–1250.
- Conn, P. B., D. S. Johnson, J. M. V. Hoef, M. B. Hooten, J. M. London, and P. L. Boveng. 2015. Using spatiotemporal statistical models to estimate animal abundance and infer ecological dynamics from survey counts. *Ecological Monographs* 85:235–252.
- Cressie, N., and C. K. Wikle. 2011. Statistics for spatio-temporal data. John Wiley & Sons, Hoboken, New Jersey, USA.
- Davis, R. A., and T. Mikosch. 2008. Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Processes and their Applications* 118:560–584.
- Davison, A. C., and M. M. Gholamrezaee. 2012. Geostatistics of extremes. *Proceedings of the Royal Society A* 468:581–608.
- Davison, A. C., S. A. Padoan, and M. Ribatet. 2012. Statistical modeling of spatial extremes. *Statistical Science* 27:161–186.
- Everitt, B. S. 1996. An introduction to finite mixture distributions. *Statistical Methods in Medical Research* 5:107–127.
- Fey, S. B., A. M. Siepielski, S. Nusslé, K. Cervantes-Yoshida, J. L. Hwan, E. R. Huber, M. J. Fey, A. Catenazzi, and S. M. Carlson. 2015. Recent shifts in the occurrence, cause, and magnitude of animal mass mortality events. *Proceedings of the National Academy of Sciences USA* 112:1083–1088.
- Finley, A. O., S. Banerjee, and B. P. Carlin. 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. *Journal of Statistical Software* 19:1–24.
- Finley, A. O., H. Sang, S. Banerjee, and A. E. Gelfand. 2009. Improving the performance of predictive process modeling for large datasets. *Computer Statistics and Data Analysis* 53:2873–2884.
- Harvell, C. D., C. E. Mitchell, J. R. Ward, S. Altizer, A. P. Dobson, R. S. Ostfeld, and M. D. Samuel. 2002. Climate warming and disease risks for terrestrial and marine biota. *Science* 296:2158–2162.
- Illian, J. B., S. Martino, S. H. Sørbye, J. B. Gallego- Fernández, M. Zunzunegui, M. P. Esquivias, and J. M. J. Travis. 2013. Fitting complex ecological point process models with integrated nested Laplace approximation. *Methods in Ecology and Evolution* 4:305–315.
- Juárez, M. A., and M. F. J. Steel. 2010. Model-based clustering of non-Gaussian panel data based on skew-t distributions. *Journal of Business and Economic Statistics* 28:52–66.
- Latimer, A. M., S. Banerjee, H. Jr Sang, E. S. Mosher, and J. A. Jr Silander. 2009. Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. *Ecology Letters* 12:144–154.
- Lindgren, F., H. Rue, and J. Lindström. 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.
- Ono, K., A. O. Shelton, E. J. Ward, J. T. Thorson, B. E. Feist, and R. Hilborn. 2016. Space-time investigation of the effects of fishing on fish populations. *Ecological Applications* 26:392–406.
- Reynolds, A. P., G. Richards, B. de la Iglesia, and V. J. Rayward-Smith. 2006. Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modeling and Algorithms* 5:475–504.
- Rindorf, A., and P. Lewy. 2006. Warm, windy winters drive cod north and homing of spawners keeps them there. *Journal of Applied Ecology* 43:445–453.
- Roislien, J., and H. Omre. 2007. T-distributed random fields: a parametric model for heavy-tailed well-log data. *Mathematical Geology* 38:821–849.
- Rose, G. A., and D. W. Kulka. 1999. Hyperaggregation of fish and fisheries: How catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. *Canadian Journal of Fisheries and Aquatic Sciences* 56:118–127.
- Rue, H., S. Martino, and N. Chopin. 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.
- Shelton, A. O., J. T. Thorson, E. J. Ward, and B. E. Feist. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. *Canadian Journal of Fisheries and Aquatic Sciences* 71:1655–1666.
- Sigrist, F., H. R. Künsch, and W. A. Stahel. 2015. Spate: an R package for spatio-temporal modeling with a stochastic advection-diffusion process. *Journal of Statistical Software* 63:1–23.
- Smale, D. A., and T. Wernberg. 2013. Extreme climatic event drives range contraction of a habitat-forming species. *Proceedings of the Royal Society B* 280:20122829.
- Stan Development Team. 2016a. RStan: the R Interface to Stan. R package version 2.14.1.
- Stan Development Team. 2016b. Stan Modeling Language User's Guide and Reference Manual, Version 2.14.0.
- Taleb, N. N. 2007. The black swan: the impact of the highly improbable. Random House, New York, New York, USA.
- Thorson, J. T., and L. A. K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* 74:1311–1321.
- Thorson, J. T., I. J. Stewart, A. E. Punt, and J. M. Jech. 2011. Accounting for fish shoals in single- and multi-species survey data using mixture distribution models. *Canadian Journal of Fisheries and Aquatic Sciences* 68:1681–1693.
- Thorson, J. T., H. J. Skaug, K. Kristensen, A. O. Shelton, E. J. Ward, J. H. Harms, and J. A. Benante. 2015. The importance of spatial models for estimating the strength of density dependence. *Ecology* 96:1202–1212.
- USDA Forest Service. 2017. Insect and Disease Survey (IDS) Database. <https://foresthealth.fs.usda.gov/>

- Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27:1413–1432.
- Ver Hoef, J. M., E. E. Peterson, M. B. Hooten, E. M. Hanks, and M.-J. Fortin. 2017. Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs* 88:36–59.
- Ward, E. J., R. Hilborn, R. G. Towell, and L. Gerber. 2007. A state-space mixture approach for estimating catastrophic events in time series data. *Canadian Journal of Fisheries and Aquatic Sciences* 64:899–910.

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/ecy.2403/supinfo>

#### DATA AVAILABILITY

All code and data associated with this paper are archived at <https://doi.org/10.5281/zenodo.1240567>. The R package glmmfields is available at <https://cran.r-project.org/package=glmmfields>.