Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures

Emeric Thibaud (Colorado State University, CSU),

Juha Aalto (Finnish Meteorological Institute),

Daniel S. Cooley (CSU),

Anthony C. Davison (Ecole Polytechnique Fédérale de Lausanne) &

Juha Heikkinen (Luke)

11th French-Danish Workshop on Spatial Statistics and Image Analysis in Biology (SSIAB 11),

Rennes, Brittany, France, May 25-27, 2016

http://www.lebesgue.fr/content/sem2016-SSIAB2016

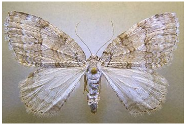
Luke

NATURAL RESOURCES INSTITUTE FINLAND

Motivation

Periodic outbreaks of insect populations can cause massive forest damages. Between 2002 and 2008 winter moth *Operophtera brumata* and autumnal moth *Epirrita autumnata* caused severe defoliation of mountain birch in an area of 10,000 km^2 (black) in northern Fennoscandia (Jepsen et al. 2009).

O. brumata



E. autumnata

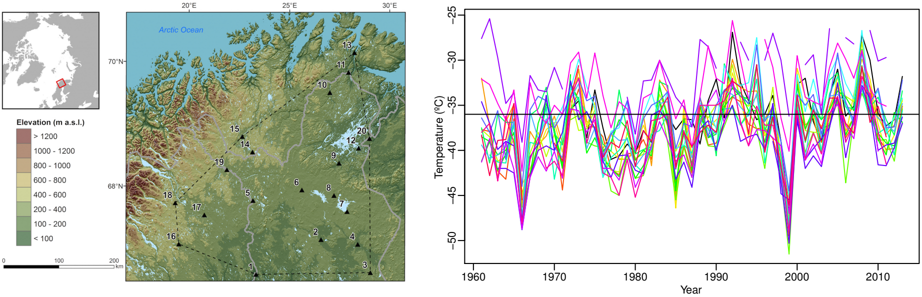
Ecology

Eggs of these moth overwinter. The next life stage, *larvae* feed on birch foliage and are the cause of the damages.

Quite sharp critical temperature limits of about -36°C for egg survival have been found (e.g., Ammunét et al. 2012). For an outbreak to develop, at least three winters with minimum temperature above the critical limit are needed (Virtanen et al. 1998).

Aim of study reported in this talk

To model the spatio-temporal pattern of winter minimum temperatures in Northern Fennoscandia using daily minimums from 20 weather stations with a relatively long series of relatively complete records.



- in order to assess spatial variation in the risk of pest insect outbreaks (or in temperatures preventing it), and
- to quantify the effects of predicted climate change.

More specific aim for future work

High-resolution spatial interpolation of winter minimum temperatures in a given year

- utilizing the developed spatio-temporal model with high-resolution covariate data (such as relative elevations) and
- conditioning on the target year's observations at the weather stations.

Comparison to observed boundaries of historical outbreaks is expected to yield independent validation of the interpolations.

Localized predictions of outbreak risk or potential spread of outbreak region obtained during the preceding winter would be valuable in implementing appropriate control strategies to minimize the damage, e.g., restricting the browsing pressure by reindeer in high-risk regions (den Herder and Niemelä 2003).

This far

Thibaud et al. (2015) developed a hierarchical spatio-temporal model based on the Brown–Resnick process for spatial extremes and a novel Bayesian approach to its parameter estimation (this talk).

Spatial interpolations for individual years (future work) with conditional simulations (Dombry et al. 2013) of the fitted model.

An essential criterion for a suitable model is that its individual realizations (predicted minimum temperature surfaces for single years) have a realistic (strong) spatial correlation structure:

- Winters tend to be cold (like 1999) or warm (2007) simultaneously at all stations and
- often the winter minimum temperatures at several stations occur on the same day or only few days apart, suggesting that they correspond to the same event.

Marginal distributions

Negated winter minimum temperature $-Y_{ij}$ for year i at weather station j was modeled as a realization from *generalized extreme-value* (GEV) distribution:

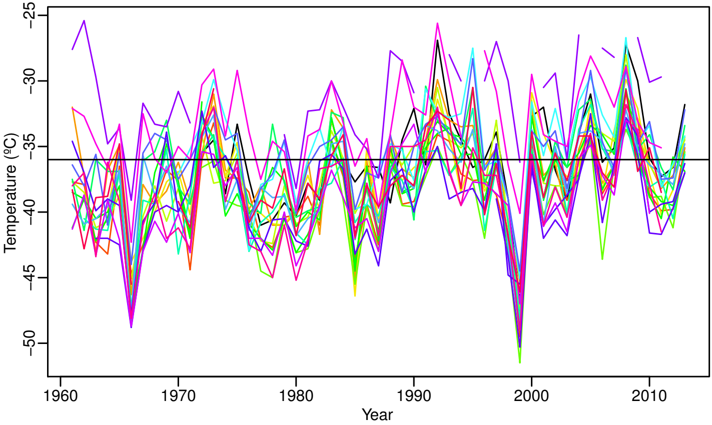
$$F(-y) = Pr(-Y_{ij} \leq -y) = \exp\left[- \left\{ 1 + \xi_{ij} \left(\frac{y - \mu_{ij}}{\sigma_{ij}} \right) \right\}_+^{-1/\xi_{ij}} \right],$$

where $a_+ = \max(a, 0)$, $\mu_{ij} \in \mathbb{R}$, $\sigma_{ij} > 0$, and $\xi_{ij} \in \mathbb{R}$.

- Three cases: Weibull ($\xi < 0$), Gumbel ($\xi = 0$), and Fréchet ($\xi > 0$) distributions.
- The case $\xi = 0$ is defined as the limit for $\xi \to 0$.
- GEV distribution is a sensible model for block maxima $-Y$.

Temporal trend

Exploratory analysis revealed a significant *temporal* trend: winter minimum temperatures have generally increased during the 53 years, although year-to-year variation is relatively large.



Spatio-temporal variation

Fitting GEV distributions at each station j separately with a linear temporal trend in *location parameters* μ_{ij} indicated that

- a flexible model was needed to capture the *spatial* variation in μ_{ij}, but
- the estimates of the slope of the temporal trend and those of the *scale* and *shape parameters* σ and ξ did not vary significantly between stations.

General year-to-year variation will be reflected in $\hat{\sigma}$, because $Y_{ij} - \mu_{ij}$ and $Y_{ij'} - \mu_{ij'}$ will be strongly correlated for all pairs j, j' of stations in the spatial model of residuals to be specified in a moment.

Space-time model for location parameters

$$\mu(\boldsymbol{s}_j, t_i) \sim GP\left\{\boldsymbol{X}(\boldsymbol{s}_j)\boldsymbol{\beta}, \tau^2\rho\right\} + \alpha t_i,$$

where

- $\boldsymbol{s}_j$ is the location of station j,
- t_i is the number of days / 365 from Jan 1, 2000 to the day of occurrence of winter minimum y_{ij},
- $\boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ is the spatial trend surface term using *coordinates*, *absolute elevations*, *relative elevations*, *proximity to the Arctic Ocean*, and *lake cover index* as covariates (Aalto et al. 2014), and
- $GP\{x(\boldsymbol{s}), \tau^2\rho\}$ denotes a Gaussian spatial process with mean $x(\boldsymbol{s})$, variance τ^2, and correlation function $\rho(\boldsymbol{h}) = \exp(-\|\boldsymbol{h}\|/\delta)$.

Scale and shape parameters σ_{ij} and ξ_{ij} were modeled as constant over both time and space.

Spatial correlation

Marginal distribution of appropriately standardized residuals

$$Z_{ij} = \{1 + \xi(Y_{ij} - \mu_{ij})/\sigma\}_{+}^{1/\xi}$$

is unit Fréchet, i.e., GEV with $\mu = \sigma = \xi = 1$.

According to exploratory analyses, residuals Z_{ij} and $Z_{ij'}$ are strongly dependent for all pair j, j' of stations.

Brown–Resnick processes (Kabluchko et al. 2009) provides a suitable max-stable model for such asymptotic dependence.

Given unit Fréchet marginals, the process is fully specified by a variogram $2\gamma(\boldsymbol{h})$ (more details in the following slides).

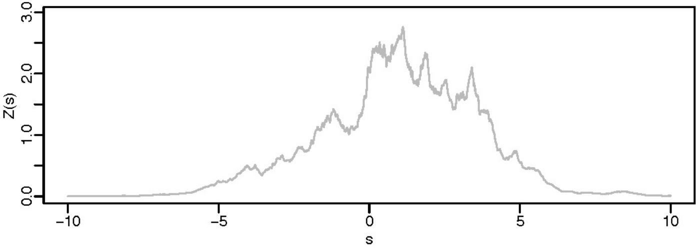
We used the stable variogram $2\gamma(\boldsymbol{h}) = 2(\|\boldsymbol{h}\|/\lambda)^{\kappa}$, $\lambda > 0$, $\kappa \in (0, 2]$.

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k\geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



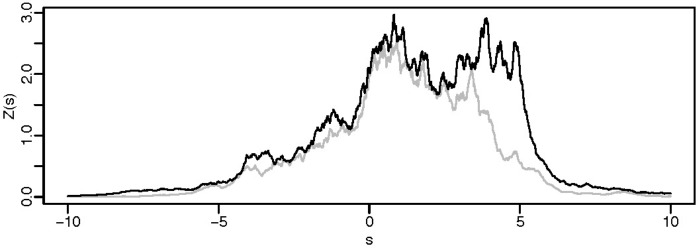
Gray: $W_1(\boldsymbol{s})$

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k \geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



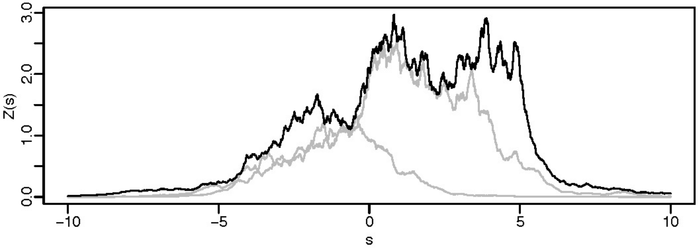
Gray: $W_k(\boldsymbol{s})$, $k = 1, 2$. Black: $\tilde{Z}(\boldsymbol{s}) = \max_{k=1,2} W_k(\boldsymbol{s})$

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k \geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



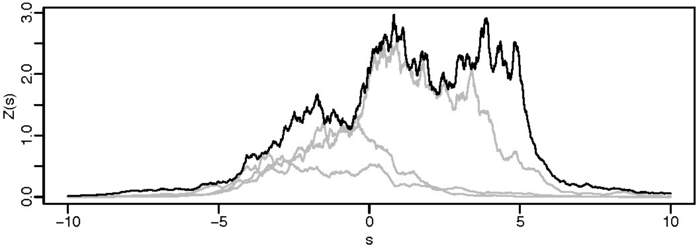
Gray: $W_k(\boldsymbol{s})$, $k = 1, 2, 3$. Black: $\tilde{Z}(\boldsymbol{s}) = \max_{k=1,\ldots,3} W_k(\boldsymbol{s})$

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k\geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



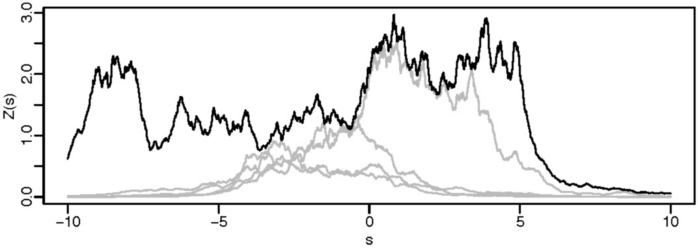
Gray: $W_k(\boldsymbol{s})$, $k = 1, \ldots, 4$. Black: $\tilde{Z}(\boldsymbol{s}) = \max_{k=1,\ldots,4} W_k(\boldsymbol{s})$

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k\geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



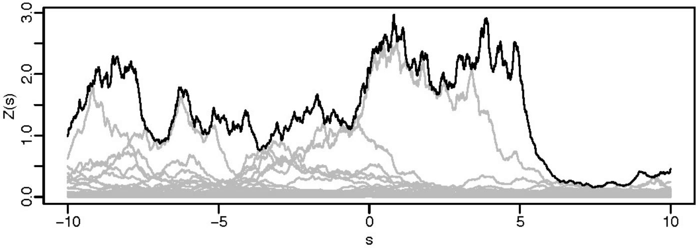
Gray: $W_k(\boldsymbol{s})$, $k = 1, \ldots, 5$. Black: $\tilde{Z}(\boldsymbol{s}) = \max_{k=1,\ldots,5} W_k(\boldsymbol{s})$

Brown–Resnick process

$$Z(\boldsymbol{s}) = \max_{k\geq 1} W_k(\boldsymbol{s}), \quad W_k(\boldsymbol{s}) = \exp\{\varepsilon_k(\boldsymbol{s}) - \gamma(\boldsymbol{s})\}/Q_k,$$

where

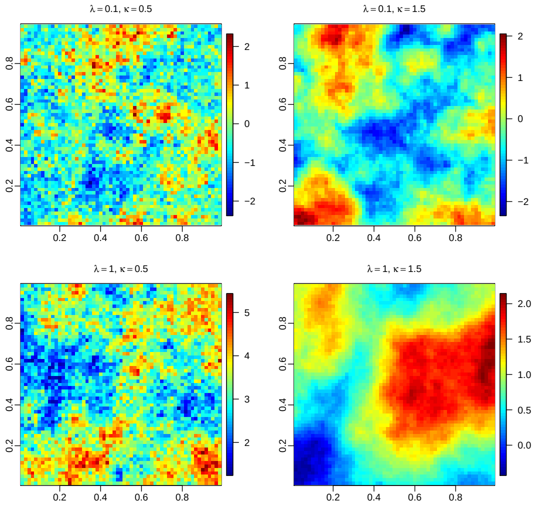
- ε_k: iid intrinsically stationary *Gaussian processes with variogram* $2\gamma(\boldsymbol{h})$ and $\varepsilon_k(\boldsymbol{0}) = 0$;
- $Q_1 < Q_2 < \cdots$: points of a unit rate Poisson process on $\mathbb{R}_+$.



Gray: $W_k(\boldsymbol{s})$, $k = 1, \ldots, 3000$. Black: $\tilde{Z}(\boldsymbol{s}) = \max_{k=1,\ldots,3000} W_k(\boldsymbol{s}) \approx Z(\boldsymbol{s})$

Realizations from Brown–Resnick process

Gumbel marginals and variogram $2\gamma(\boldsymbol{h}) = 2(\|\boldsymbol{h}\|/\lambda)^{\kappa}$.



Likelihood function

Contribution of one year's residuals $Z_i = \{Z_{i1}, Z_{i2} \dots, Z_{iD}\}$ to the likelihood function for the Brown–Resnick process is of the form

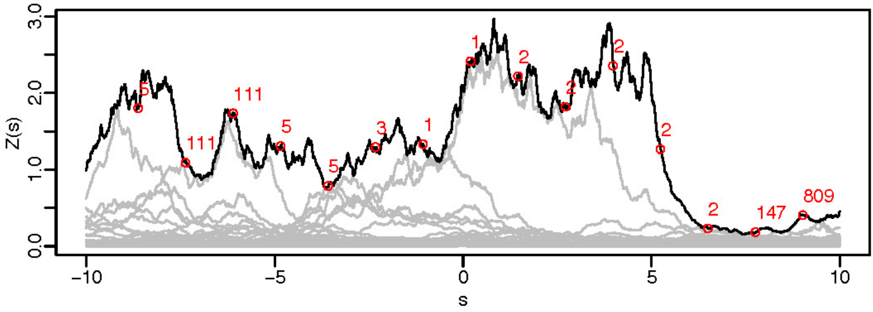
$$\sum_{\Pi \in \mathcal{P}} \left[\exp\{-V(Z_i)\} \prod_{\pi_m \in \Pi} \left\{ -V_{\pi_m}(Z_i) \right\} \right], \tag{1}$$

where the sum is over all possible partitions of the set $\{1, \dots, D\}$ of stations.

For $D = 20$, $card(\mathcal{P}) \approx 10^{13}$: the likelihood is intractable.

Each partition Π in (1) represents one possibility to generate $\boldsymbol{Z}$ from the latent $\boldsymbol{W}_k$'s:

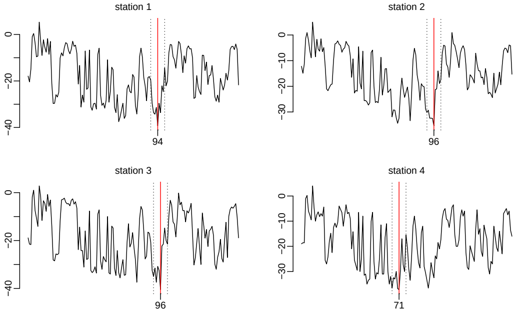
- $Z_j = \max_{k \geq 1} \boldsymbol{W}_{kj}$
- The term with partition $\Pi = \{\pi_1, \pi_2, \ldots, \pi_M\}$ covers cases, where both Z_j and $Z_{j'}$ originate from the same $\boldsymbol{W}_k$ whenever $\{j, j'\} \subset \pi_m$ for some $m \in 1, \ldots, M$.



Here, with $n = 15$ sites (numbered 1 to 15 from left to right),
$\Pi = \{(1,4,5),(2,3),(6),(7,8),(9,10,11,12,13),(14),(15)\}$.

M2: Stephenson and Tawn (2005) model

- Condition on fixed partitions Π_i pre-estimated using *declustering* based on dates of the extreme events ⇒ (1) reduces to one term.
- In our application we assumed that, for each year, groups of stations, where subsequent minima are separated by at most 5 days, belong to the same cluster π_m.



Example: here the partition is $\{(1,2,3),(4)\}$

M3: Random partition model

Fixed declustering may lead to bias: true distribution of *latent* partitions can be very different.

The new approach presented by Thibaud et al. (2015) uses *data augmentation principle*:

- introduce the *unknown partition* Π into the Bayesian hierarchical model, and
- impute its values using an *MCMC algorithm* (Dombry et al. 2013).

Comparison of models

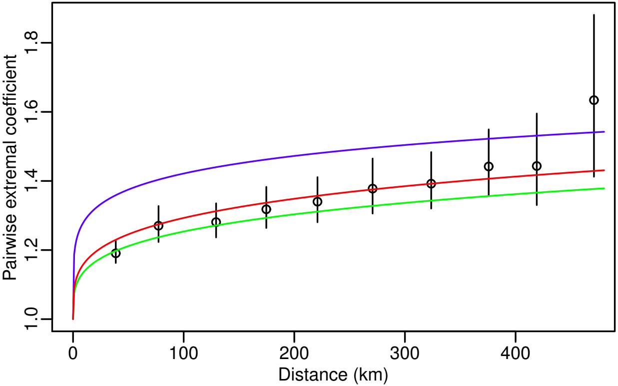
- M1 Trend surface model (the standard approach).
- M2 Stephenson and Tawn (2005) model based on the declustering.
- M3 Thibaud et al. (2015) model with random partitions.

Parameter estimates (mean of posterior distributions) and 95% confidence/credible intervals from M1 (top), M2 (middle), and M3 (bottom). The first column is the estimate for the mean value of the location parameter $\mu_j = \mu(\boldsymbol{s}_j, 0)$ over the 20 stations.

$-\bar{\hat{\mu}}_j$	$-\hat{\alpha}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\lambda}$	$\hat{\kappa}$
$-34.9_{(-35.7,-34.1)}$	$0.07_{(0.03,0.10)}$	$3.4_{(3.0,3.7)}$	$-0.15_{(-0.18,-0.11)}$	$368_{(31,705)}$	$0.37_{(0.30,0.44)}$
$-34.6_{(-35.6,-33.7)}$	$0.09_{(0.05,0.13)}$	$4.0_{(3.6,4.6)}$	$-0.11_{(-0.15,-0.06)}$	$1811_{(977,3615)}$	$0.54_{(0.46,0.63)}$
$-34.9_{(-35.9,-33.9)}$	$0.06_{(0.01,0.11)}$	$3.5_{(3.0,4.1)}$	$-0.10_{(-0.14,-0.06)}$	$1086_{(462,3065)}$	$0.53_{(0.43,0.64)}$

M2 vs. M3: Larger $\hat{\sigma}$ and $\hat{\lambda}$. Model diagnostics support M3.

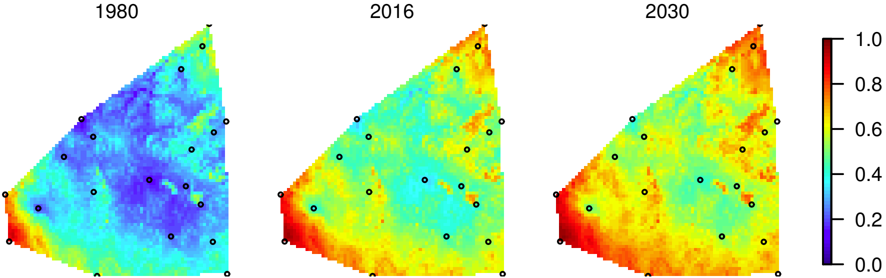
Example of diagnostics: etremal coefficients



Binned estimates of empirical pairwise extremal coefficients, with their 95% confidence intervals. Curves are from the fitted Brown–Resnick models: M1 (blue), M2 (green), and M3 (red).

M3: Predictions for annual minimum temperatures

The model predicts an increase of 0.6°C per decade.



(Unconditional) probability that the annual minimum temperature exceeds -36°C for 1980, 2016 and 2030.

Conclusion and further...

A flexible framework for spatio-temporal modelling of extremes, but complicated and computationally intensive: cannot be used for large datasets.

Linear temporal trend can be replaced by climate model predictions to enable more realistic forecasts.

Conditional simulations and validation against spatial extent of observed outbreaks remains to be done.

Another application of interest to us: Spatial predictions of the occurrence of frost during flowering time of wild berries validated against observed frost damages / yields.

Aalto, J., P. C. le Roux, and M. Luoto. 2014. The meso-scale drivers of temperature extremes in high-latitude Fennoscandia. Climate Dynamics **42**:237–252.

Ammunét, T., T. Kaukoranta, K. Saikkonen, T. Repo, and T. Klemola. 2012. Invading and resident defoliators in a changing climate: cold tolerance and predictions concerning extreme winter cold as a range-limiting factor. Ecological Entomology **37**:212–220.

den Herder, M., and P. Niemelä. 2003. Effects of reindeer on the re-establishment of *Betula pubescens* subsp. *czerepanovii* and *Salix phylicifolia* in a subarctic meadow. Rangifer **23**:3–13.

Dombry, C., F. Éyi-Minko, and M. Ribatet. 2013. Conditional simulation of max-stable processes. Biometrika **100**:111–124.

Jepsen, J., S. Hagen, K. Høgda, R. Ims, S. Karlsen, H. Tømmervik, and N. Yoccoz. 2009. Monitoring the spatio-temporal dynamics of geometrid moth outbreaks in birch forest using MODIS-NDVI data. Remote Sensing of Environment **113**:1939–1947.

Kabluchko, Z., M. Schlather, and L. de Haan. 2009. Stationary max-stable fields associated to negative definite functions. The Annals of Probability **37**:2042–2065.

Stephenson, A., and J. Tawn. 2005. Exploiting occurrence times in likelihood inference for componentwise maxima. Biometrika **92**:213–227.

Thibaud, E., J. Aalto, D. S. Cooley, A. C. Davison, and J. Heikkinen. 2015. Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures. Technical report, arXiv:1506.07836.

Virtanen, T., S. Neuvonen, and A. Nikula. 1998. Modelling topoclimatic patterns of egg mortality of *Epirrita autumnata* (Lepidoptera: Geometridae) with a geographical information system: predictions for current climate and warmer climate scenarios. Journal of Applied Ecology 35 **35**:311–322.

Thank you!

NATURAL RESOURCES INSTITUTE FINLAND