Modelling disease spread when the populations at risk are poorly mapped

Stephen Catterall[1], Glenn Marion[1] and Thibaud Porphyre[2]

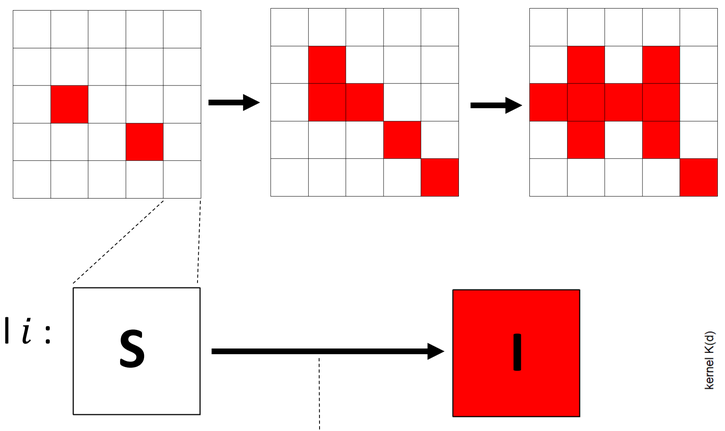
1 Biomathematics and Statistics Scotland, Edinburgh, United Kingdom

2 Laboratoire de Biométrie et Biologie Évolutive, CNRS, VetAgro Sup, Marcy l'Étoile, France

1: the challenge

- Disease outbreak data typically consist only of case reports with times/locations
- However, most outbreak data analysis methods require knowledge of the spatial distribution of the population at risk
- **Problem**: there is often much uncertainty in this distribution of potential hosts because
  - Data simply do not exist e.g. wildlife populations
  - Data are not publicly available for privacy reasons e.g. farm locations in the U.S.A.

2: iPAR, our approach

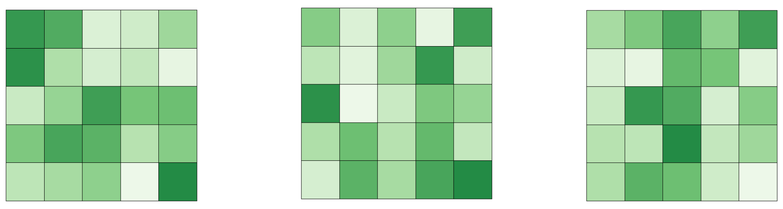


Grid cells become infected over time

force of infection on i $f_i = \varepsilon s_i + \rho \sum_j s_i t_j K(d_{ij})$

Kernel captures distance dependence in transmission

$s_i = \left(\sum_{i=1}^{L} h_{i,l}\sigma_l\right)\exp\left(\sum_{k=1}^{K} c_{i,k}\sigma'_k\right) \quad \sigma \in S^L \quad \sigma' \in \mathbb{R}^K$



Chosen spatial covariates reflect host population density and many other factors such as biosecurity and the relative susceptibility of hosts

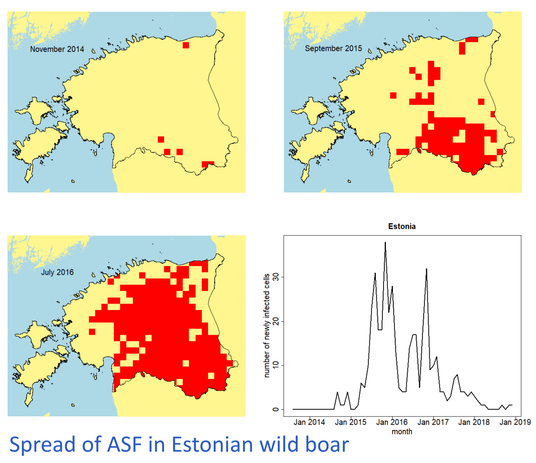
Spatial covariates $h_{i,l}$ and $c_{i,k}$ used to explain variation in susceptibility s_i
A similar approach is used to model variation in infectivity t_j

- To allow analysis of disease case data when the population at risk is unknown, we assume that the host's spatial distribution is associated with known spatial covariates e.g. land use data, climate data
- The nature of this association is informed by the outbreak data
- Parameters are estimated via Bayesian MCMC

- Land use data is typically compositional data
- The model also includes compositional parameters, for example σ belongs to S^L, which is the L dimensional simplex

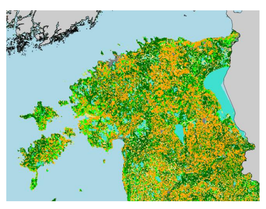
For further details see PLOS Comp Bio article
https://dx.plos.org/10.1371/journal.pcbi.1012622

3: application to African Swine Fever in Estonian wild boar

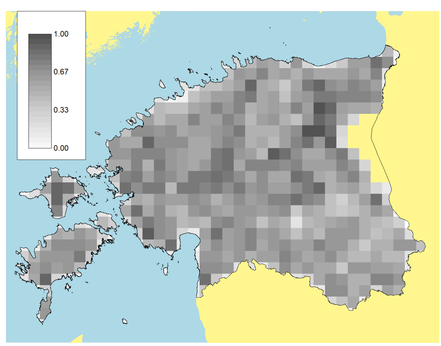


Spread of ASF in Estonian wild boar according to spatially aggregated case reports

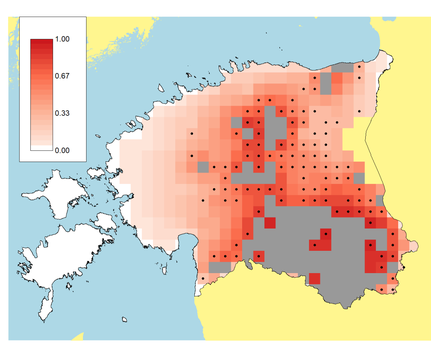
MODEL INPUTS



CORINE land use data (grey=urban, orange=arable, light green=broadleaf, dark green=conifer, turquoise=water)



Estimated susceptibility across Estonia based on case reports of ASF in wild boar



Risk map for spread of ASF in wild boar in Estonia. Grey patches are recorded as infected at the end of the early stage of the outbreak. Red colours indicate probability of infection four months beyond the end of the early stage, based on a model fitted to early-stage data. The black dots show patches that did in fact become infected within this time horizon

SELECTED MODEL OUTPUTS

4: discussion

Key innovations

- Joint inference of local susceptibility and infectivity in a data-poor setting
- Compositional parameters and data – connections to compositional data analysis
- Estimate temporal changes associated with management

Future work

- Extension to multiple species
- Integration of control measures into the simulation component of the framework
- Application to other diseases – avian influenza, chronic wasting disease?

Related questions regarding model transfer

- How can outputs from the above fitted model be used to inform risk assessments for what might happen if ASF arrives in the U.K.?

Scan to visit journal article