

Supporting Information – Appendix 1

Description of Generalized Additive Model (GAM) structure and assessment of fit

We used hierarchical generalized additive models (HGAMs, Pederson et al. 2019) to assess annual trends in the number of reported incidents (N) across villages (Vil), damage types (DTy) and species (Spp) involved. For each of these grouping factors, we built and compared three different model structures.

Model 1 assumes no variation across factor levels (i.e. individual villages, damage types or species). For a given grouping factor, the general formula for this model is:

$$N = f(month) + \zeta_{year} + \varepsilon$$

where f is a smooth function, ζ_{year} is the random effect for year and ε is a negative binomial error term.

The corresponding R code used to implement this model (package *mgcv*) is:

```
model1 <- gam(N ~ s(year, bs="re") + s(month, k=12), data=data,
              family=nb, method="REML", drop.unused.levels=F,
              offset=log(village.population.size))
```

Model 2 allows annual trend to vary independently across factor levels.

Mathematically, this is written as:

$$N = f_{grouping.factor}(month) + \zeta_{year} + \varepsilon$$

The corresponding R code is:

```
model2 <- gam(N ~ s(year, bs="re") + s(month, grouping.factor,
                                       bs="fs", k=12), data=data, family=nb, method="REML",
              drop.unused.levels=F,
```

```
offset=log(village.population.size))
```

Lastly, Model 3 also allows for variation across factor levels, but with a penalty for deviations from a global shared trend. This penalty is based on the squared distance between the factor level smooth and an average shared smooth (i.e. global shared trend). In other words, Model 3 assumes each factor level curve has a similar shape to the others (see Pedersen et al. 2019 for more details). Mathematically, this model can be represented as:

$$N = f(\text{month}) + f_{\text{grouping.factor}} + \zeta_{\text{year}} + \varepsilon$$

The corresponding R code is:

```
Model3 <- gam(N ~ s(year, bs="re") + s(month, grouping.factor,  
      bs="fs", k=12, m=2), data=data, family=nb,  
      method="REML", drop.unused.levels=F,  
      offset=log(village.population.size))
```

All models were assessed based on the Deviance Explained D , which is calculated as:

$$D = 2[l(\hat{\beta}_{max}) - l(\hat{\beta})] \phi$$

In which $l(\hat{\beta}_{max})$ is the maximized likelihood of the saturated model, $l(\hat{\beta})$ is the maximized likelihood of the fitted model, and ϕ is a scaling parameter (see Wood 2006 for more details).

References

Pedersen EJ, Miller DL, Simpson GL, Ross N. 2019. Hierarchical Generalized Additive Models in Ecology: an introduction with mgcv. *PeerJ* 7: e6876.

Wood SN. 2006. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.