Hindcasting Endemic Disease - Foot and Mouth Disease

Gustaf Rydevik

25 July 2016

Hindcasting FMD incidence in Cameroon using surveillance data with multiple diagnostics

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

Introduction

This section describes an analysis of data collected on FMD prevalence in a dataset of 1500 cattle distributed across XX herds. The goal of the data collection was to get information on the time since last FMD incursion in a herd. Various aspects of the dataset has been described and analyzed in a number of publications [1,2,3,...]. Here, we will describe how the prevalence data can be transformed to provide information on disease incidence.

Data

The data used was collected by Bronsvoort etal in Cameroon. Full details can be found in ....

There was the following diagnostic measures used on included cattle:

-Danish C-ELISA FOr non-structural protein, so will pick up any serotype in theory.

kinetics from Bronsvoort2004

-South American I-ELISA Also non-structural and will pick up any serotype

-CHEKIT kit Final non-structural (least good, "sensitivity" of ~23%)

kinetics from Bronsvoort2004

-EITB - enzyme linked immunoelectrotransfer blot Non structural, but a binary yes/no response. Americans use EITB and I-ELISA as a combined diagnostic

FMD_VNT are virus neutralisation test, where you do sequential dilutions of the virus and look whether antibodies react or not.

Higher the number, the greater the dilution with detectable virus levels. FMD has seven different serotypes, with different geographical distributions. Three VNT tests: A-serotype, O-serotype and SAT2 ("south-african territories") serotype.

PRobang is used to scrape the cells to collect viruses for cultivation. Only done on animals where we think that something is going on. That's PbP A,o,and Sat2. From herds 32 onwards, we have collected the probang. The numbers are (probably) binary classifiers based on Antigen ELISA-results.

Finally, FMDS O, A, and SAT2 are the binary classifiers based on the VNT tests.

FMD_VNT_A

FMD_VNT_O

FMD_VNT_SAT2

If an animal gets exposed, the animal will develop a particular antibody response that will last for years. The VNT results will therefore likely remain for a long time following exposure/infection. The non-structural ELISA results will disappear within 6-12 months. The probang has some sort of exponential decay of viruses/ likelihood of cultivation. So a latent probability of some sort...

An additional thing for the future is that we have recording based on the growth of hoofs and how far up there are records of old lesions.

In terms of animals, we have clinically infected, recently infected, old lesions, and healthy animals.

Monlast is the herdsmans reporting, and can be used as a validation to compare with estimated times since infection.

One of the practical questions is to go through a herd and try and age lesions. But they did run into problems with sheep flocks that they missed in 2001, since sheep tends to not have clnical signs. Using a stat method would be useful here - something to bring up in a discussion. In pig farms, at least three infection cycles before the outbreak was detected. Want to try and identify how far back was the herd infected. Being able3 to do this quite quickly would be very useful.

We can use Alexanderssen2003 as how a VNT response would look like post infection, but obviously conditional on a scaling factor to account for different units. Would go off the contact one.

Names of researchers:

Bergmann (iELISA) Bronsvoort Sorensen KJ (cELISA) Alexandersen Hamblin Brocchi Cuncliffe Dekker (review!)

People doing long-term experimental infection/post-disease follow up. Possibly Bergmann Pirbright does short, 1-month studies.

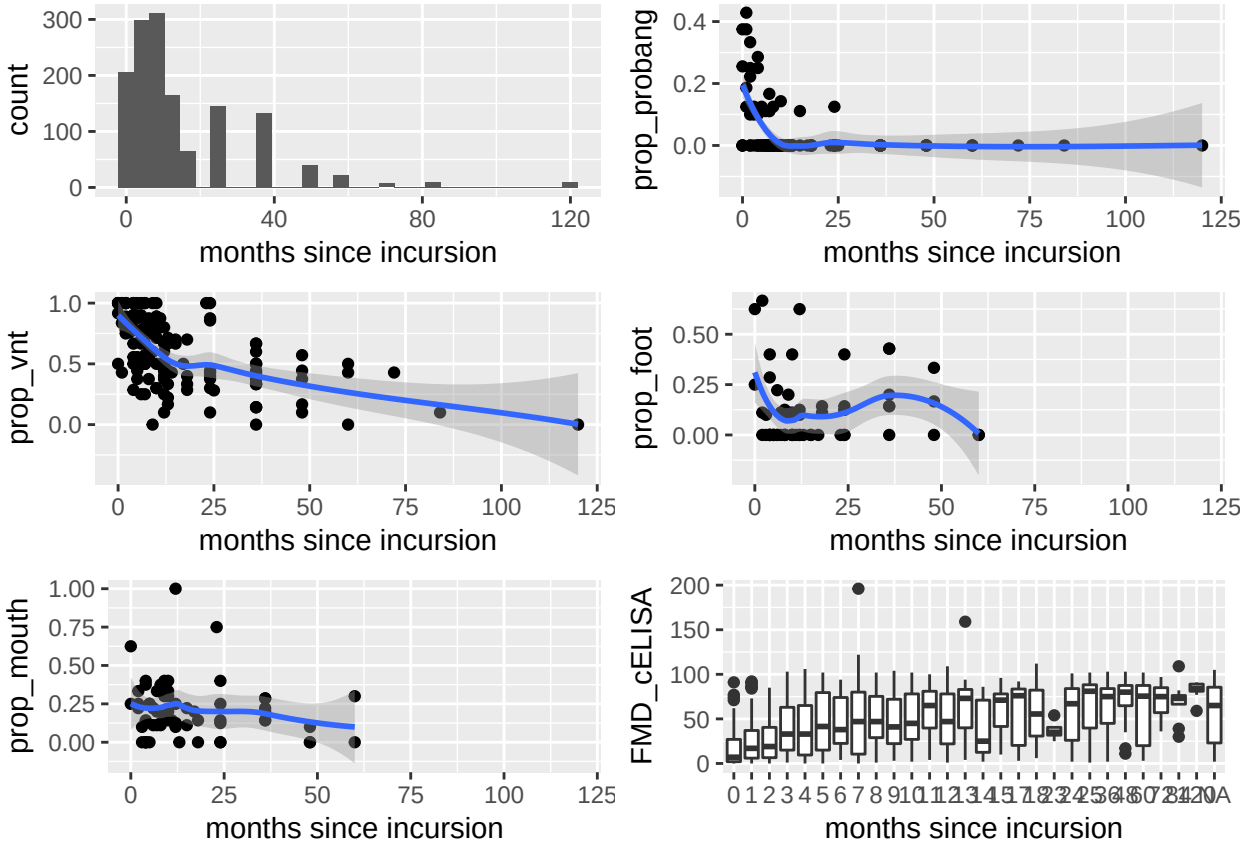
Section 1: using the cameroon data to estimate time since incursion.

```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 27 rows containing non-finite values (stat_bin).
## `geom_smooth()` using method = 'loess'
## Warning: Removed 29 rows containing non-finite values (stat_smooth).
## Warning: Removed 29 rows containing missing values (geom_point).
## `geom_smooth()` using method = 'loess'
## Warning: Removed 3 rows containing non-finite values (stat_smooth).
## Warning: Removed 3 rows containing missing values (geom_point).
## `geom_smooth()` using method = 'loess'
## Warning: Removed 87 rows containing non-finite values (stat_smooth).
## Warning: Removed 87 rows containing missing values (geom_point).
## `geom_smooth()` using method = 'loess'
## Warning: Removed 87 rows containing non-finite values (stat_smooth).
## Warning: Removed 87 rows containing missing values (geom_point).
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```



A basic model, just using the cameroon data itself, would likely look something like the following:

We have diagnostic data at the individual cow level; Probang measurements, VNT measurements, and aged lesions: $y_{i1} = Probang_i$ (which is binary) $y_{i2} = VNT_i$ (which could be treated as binary or as continous) $y_{i3} = Lesion_i$ (which is ordinal).

These are all measurements of three different latent processes. The probang results measures whether or not it was positive to successfully culture FMDV from the throat(?) of the animal, and can thus be seen as an indirect indicator of the amount of virus present in the throat. This is the first latent component, $l_{i1}(d_i) = exp(d_i, \theta_{probang})$. We here assume that the the virus declines exponentially following infection, which is a simplistic but useful approximation. [Is it?]

The VNT results is an indicator of the level of FMD-specific antibodies present in the blood of the animal. This is the second latent component, $l_{i2}(d_i) = Logistic(d_i, \theta_{VNT})$. In this case we assume that the antibody response follows a logistic growth curve.

Finally, the lesion aging attempts to measure the amount of new growth that has occured on the hoves of animals since an FMD infection (which leaves very characteristic markings at the join between legs and hooves,that then migrate downwards). $l_{i3}(d_i) = d_i * \theta_{lesions}$. We will simply model this process as a linear function of time since infection.

The full latent process can thus be written as

$P(L|T, E, \theta_L) = exp(d_i, \theta_{probang}), Logistic(d_i, \theta_{VNT}), d_i * \theta_{lesions})$

$monlast\ prior(f(incidence))$

$P(probang - positive, VNT - Positive)\ f(monlast, age, ...)$ ## Could separate, or not... $VNT \sim f(monlast)\ ELISA \sim F(monlast)$

$P(probang = 1) = Bernoulli(logit(\theta_1))\ \theta = f_1(T_{infected}) + g_1(age, age^2))$

$P(Vnt = 1) = Bernoulli(logit(\theta_2))\ \theta = f_2(T_{infected}) + g_2(age, age^2))$