Data dictionary

The file AU-FMD-epidata.csv contains epidemiological information for each infected herd in a simulated Foot and Mouth disease outbreak in Australia. The file AU-FMD-sequences.fasta contains corresponding simulated genetic sequences for each infected herd. For simplicity, we only consider one sequence per herd (say, the first detected case in that herd).

Variables available for each infected premises (IP) of foot and-mouth disease (FMD)

Herd ID - Unique identifiers herd.name and herd.id

Geocoordinates - Latitude and longitude in GDA94 coordinate reference system

Herd type - Predominant type of the herd: beef intensive, beef extensive, dairy cattle, mixed beef, feedlot, sheep, mixed sheep, pigs small, pigs large or smallholder (mixed)

Number of animals – Total in herd, for FMD susceptible species only

Day of onset - Onset of clinical signs in the earliest animal detected on this premises.

Day of diagnosis - Also assumed to be the day the genetic sample was collected

Reason diagnosed - Index premises 1^{ST}_DETECTION (where the first case was found), suspect premises (a suspected FMD case in the herd), at-risk premises (no evidence of a case, but herd considered at high risk), dangerous contact premises (known contact with 'dangerous' source e.g. another herd, through a market or suspected indirect transmission)

Day culling started - Or estimated last day infectious if not culled whilst infectious. To minimise the spread of FMD, herds are typically culled soon after a case is detected.

The spread of disease can be assumed to be along the following pathways:

• Direct contact – movement of live animals between premises,

• Market/saleyard spread – movement of live animals in and out of markets/saleyards,

• Indirect contact – movement of animal products, byproducts or fomites between herds,

• Local spread – proximity-based contact, e.g., over a boundary fence shared by adjoining premises*,

• Airborne transmission – virus excreted by animals in aerosol form that remains viable in the air.

*Defined as spread between IPs within 3 km of each other. Several possibilities include short distance aerosol spread, contamination in area near the IP e.g. infected materials on roads or common facilities