ASF CHALLENGE
MODEL DATA DOCUMENTATION
UK TEAM

Phase 1

A. Data files - wild boar model

A.1. Model inputs (folder: output-data/model input)

1. *infectionStatusMatrix.RDS*

Class: matrix

Description	Comment
Indicator matrix of infection statuses of patches by day 0	Matrix has 50 columns and 50 rows as there are 2500 patches 1 = Infected 0 = Not infected A patch is infected if at least one boar in the patch is infected.

2. *in.fence.matrix.RDS*

Class: matrix

Description	Comment
Indicator matrix for whether a patch is in the fence	1 = Within fence 0 = Outside fence

3. *patchcentres.RDS*

Class: data frame

Description: Patch coordinates in the X and Y directions. Dimension: 50 x 2

Column name	Short description	Additional comments
X	Patch coordinates in the X-direction	1 = Within fence-buffer 0 = Outside fence-buffer
Y	Patch coordinates in the Y-direction	

4. all.patch.centres.RDS

Class: data frame

Description: Information on patch location and number of boar within patches. Dimension: 2500 x 5

Column name	Short description	Additional comments
X	X coordinate of patch	1 = Within fence-buffer 0 = Outside fence-buffer
Y	Y-coordinate of patch	
region	Administrative region in which patch center is located	
number.of.boar	Number of boar within patch	
id	Unique patch ID	In the range [1, 2500] as there are 2500 patches

5. nBoarsMatrix.RDS

Class: matrix

Description	Comment
Matrix containing number of boar in each patch	Matrix has 50 columns and 50 rows

6. observed.positive.RDS

Class: data frame

Description: Data on daily number of detected positive cases, from day 1 to day 50. This is only used directly in the parameter estimation process for computing summary statistics of the observed data.

Column	Short description	Additional comments
day	Day on which observation was made	
count	Number of detected infected boar on *day*	Detection is by all methods

7. observed.positive.locations.RDS

Class: list

Description: List containing coordinates of detected positive boar *by* day *X*, where $X = \{1, 2, \ldots, 50\}$. The list has 50 elements (the first element corresponds to day 1 and the last element corresponds to day 50). Each element of the list is a data frame with columns as described below.

Column	Short description	Additional comments
X	X coordinate of infected patch	
Y	Y coordinate of infected patch	

8. removals.RDS

Class: data frame

Description: Daily data on boar removals within the area considered, from day 1 to day 50 (See supplementary material for details on area considered.)

Column name	Short description	Additional comments
in.zone	Daily number of boar removed within zone by all methods	Removal methods considered are

		hunting, active search (AS) and passive surveillance (PS).
in.zone.found	Daily number of boar found by AS or PS within zone	
in.zone.hunted	Daily number of boar hunted within zone	From day 1 to day 59, this was computed by applying the estimate for the percentage of hunted boar that were tested.
out.zone	Daily number of boar removed outside zone by all methods	
out.zone.found	Daily number of boar found by AS or PS outside zone	
out.zone.hunted	Daily number of boar hunted outside zone	
total	Total number of boar removed by all methods	

A.2. Model output (folder: output-data/model output)

A list (*sample_model_ouptut_phase1.RDS) containing the elements:**

summarized.res: Data frame of daily and cumulative counts of detected positive boar.

Column	**Short description**	**Additional comments**
day	Days for which simulation was performed	For parameter estimation, day ranged from 1 to 50.
cumulative.cases.patches	Cumulative number of infected boar by *day* as determined by the model	
cumulative.cases.wb	Cumulative number of detected infected boar by	For an infected patch, the expected number of infected boar in the patch

	day as determined by the model	was calculated as the product of the number of boar in the patch on *day*, the fraction of positive boar in the patch and the patch-level detection rate for positive boar (see Table 5 of our Model description document for details on these parameterssample).
new.cases.wb	Daily number of detected infected boar at *day* as determined by the model	

***status.matrices*:** List with length equal to maximum time, with each list element corresponding to a day. Each list element is an indicator matrix showing which patches are infected at the corresponding day: 1 for infected and 0 if not infected.

***coord.infected*:** List of length equal to the last day of simulation, with each list element corresponding to a day. Each list element is a data frame of XY coordinates of infected patches by the corresponding day.

*Note that *sample_model_ouptut_phase1.RDS* is only intended to serve as a sample which reflects the structure of the output of the model. The output values should therefore not be regarded as typical model outcomes.

A.3. Parameter estimates (folder: output-data/parameter estimates)

The parameter values accepted in the Approximate Bayesian Computation (ABC) estimation, along with the output from the corresponding simulations, were stored in a list.

The list (*results_gen_4),* is constituted of the following elements.

estims: Data frame with dimensions N x 4, where N is the number of accepted particles (parameter values which produced simulations sufficiently close to the observations, where closeness is here judged by user-supplied tolerance values). In Phase 1, $N = 500$. The columns of the data frame are described below.

Column	Short description
V1	Estimates for the transmission rate parameter.
V2	Estimates for the parameter representing the spatial scale of transmission.
V3	Estimates for the parameter representing the patch-level positivity rate before day 28 (after the index case in domestic herds).
V4	Estimates for the parameter representing the patch-level positivity rate on and after day 38 (after the index case in domestic herds).

mat: List with length equal to *N*; each list element corresponds to an accepted particle. Each list element is a list of 50 indicator matrices, with the matrix in the first position corresponding to day 1 and the matrix in the last position corresponding to day 50. These matrices show which patches, according to the model, are infected by each day: the *(i, j)*th entry is 1 if patch *(i, j)* is infected and 0 if it is not infected.

A.4. Predictions (folder: output-data/predictions)

The parameter estimates (with corresponding simulation outputs) were run through a prediction model to obtain predictions of case counts (total and detected) and infection locations up to day 78 under the increased hunting pressure and scenario and normal hunting pressure scenarios

The data files containing the predictions are:

File	Description	Details
forward.run.RDS	Predictions from day 51 to day 78 assuming no fence and normal hunting pressure	A list with the elements: ***status:*** List with length *N*. Each list element corresponds to one accepted simulation and is a list of indicator matrices for the infection status of patches, as predicted by the model. Each sublist holds 28 elements: the first and last elements hold matrices

		corresponding to day 51 and by day 78 respectively. These elements have the same structure as *status.matrices* (see Section A.2.). ***coord***: List with length *N*. Each list element corresponds to one accepted simulation and is a list of locations of infections in patches, as predicted by the model. Each sublist holds 28 elements: the first and last elements hold XY coordinates of infected patch locations by day 51 and by day 78 respectively. These elements have the same structure as *coord.infected* (see Section A.2.). ***counts:*** List with length *N*. Each list element corresponds to one accepted simulation and is a data frame of predicted infection counts having the same structure as *summarized.res* (see Section A.2.).
forward.run.fence.RDS	Predictions from day 51 to day 78 assuming a fence and normal hunting pressure within the fence	As above
forward.run.fence.increased.pressure.RDS	Predictions from day 51 to day 78 assuming a fence and increased hunting pressure within the fence (and normal hunting pressure outside the fence, as in the above 2 scenarios)	As above

B. Data files – pig herd model

B.1. Pig movements input

1. ***final_ERGM.Rdata (folder Moves_ERGM)***

   Object: netmodel.3
   Class: ergm

Description: Selected (final) model used to simulate pig movements between herds. See "?ergm" for the elements of an object of class ergm. Note that the method summary.ergm returns a summary of the relevant parts of the ergm object in concise summary format.

2. ***posnegbinom.Rdata (folder Moves_GLM)***

   Object: mod2
   Class: vglm

Description: Selected model used to simulate the number of pigs in shipments. See "?vglm" for the elements of an object of class vglm.

B.2. Pig movements output

3. ***mov_pred_i.Rdata with $1 \leq i \leq 500$ (folder Moves_Simul/sim)***

   Object: mov.pred
   Class: data.frame

Description: Simulated pig movements from day 51 up to day 80.

Column name	Short description	Additional comments
source	Unique identification number of the pig site that moved the pigs	Same as in *moves_Players_day_50.csv*
dest	Unique identification number of the pig site that received the pigs	
source.type	Production type of source herd	
dest.type	Production type of dest herd	
date	Day of the shipment	
source.size	Number of pigs in the source herd	Same as in *herds_day_50.csv*
dest.size	Number of pigs in the dest herd	Same as in *herds_day_50.csv*
source.com	Activity of the source herd (commercial/backyard)	1 for commercial sites; 0 for backyard sites
dest.com	Activity of the dest herd (commercial/backyard)	1 for commercial sites; 0 for backyard

		sites
source.out	Whether the pigs have access to an outdoor area or not in the source herd	1 for access to outdoor; 0 otherwise
dest.out	Whether the pigs have access to an outdoor area or not in the dest herd	1 for access to outdoor; 0 otherwise
source.multi	Unique identification number of the multisite farms to which the source herd belongs	0 for herds that do not belong to any multi-site farms; $i>0$ for herds that belong to a multisite farm
dest.multi	Unique identification number of the multisite farms to which the dest herd belongs	0 for herds that do not belong to any multi-site farms; $i>0$ for herds that belong to a multisite farm
multi	Whether the source and dest herds belong to the same multisite	1 if source and dest herds belong to the same multi-site farm; 0 if they do not belong to the same multi-site farm or if they do not belong to any multi-site farms
qty	Total number of pigs of the shipment	Same as in *moves_Players_day_50.csv*

4. ***mov_Se_i.Rdata with $1 \leq i \leq 500$ (folder Moves_Simul/Se)***

   Object: Se
   Class: numeric

Description: A vector containing the sensitivity of the simulated movement network (proportion of the existing edges between pairs of herds in the observed movement data (last 30 days not used in model selection and fitting) reproduced by the simulated movement network).

5. ***mov_Sp_i.Rdata with $1 \leq i \leq 500$ (folder Moves_Simul/Sp)***

   Object: Sp
   Class: numeric

Description: A vector containing the specificity of the simulated movement network (proportion of the possible pairs of herds that were not connected in the observed movement data (last 30 days not used in model selection and fitting) reproduced by the simulated movement network).

B.3. Pig herd model input

Folder Moves_Simul/sim:

- *mov_pred_i.Rdata with 1 ≤ i ≤ 500 (folder Moves_Simul/sim):* see section B.2. Only columns "*date*", "*source*", "*source.type*", "*dest*", "*dest.type*" and "*qty*" (same as in moves.csv) are used. In phase one, only moves between day 51 and day 78 were used.

Folder Model_Predict/input:

- *patchcentres.RDS:* see section A.1.
- *nBoarsMatrix.RDS*: see section A.1.
- *results_gen_4.RDS:* see section A.3.
- *forward.run.RDS:* see section A.4.
- *forward.run.fence.RDS:* see section A.4.
- *forward.run.fence.increased.pressure.RDS:* see section A.4.

6. ***forward.pressure.RDS (folder Moves_Predict/input)***

   Object: final.pressure
   Class: list

Description: List of length 500, with each component containing a matrix of dimensions of 4533 * 78 with the infectious pressures exerted by wild boar on each herd (4533) and at each time step in the wild boar model (78) for the scenario without the fence.

7. ***forward.pressure.fence.RDS (folder Moves_Predict/input)***

   Object: final.pressure
   Class: list

Description: List of length 500, with each component containing a matrix of dimensions of 4533 * 78 with the infectious pressures exerted by wild boar on each herd (4533) and at each time step in the wild boar model (78) for the scenario with the fence and normal hunting pressure.

8. ***forward.pressure.fence.increased.hunting.RDS (folder Moves_Predict/input)***

   Object: final.pressure
   Class: list

Description: List of length 500, with each component containing a matrix of dimensions of 4533 * 78 with the infectious pressures exerted by wild boar on each herd (4533) and at each time step in the wild boar model (78) for the scenario with the fence and increased hunting pressure.

B.4. Pig herd model output

9. ***simul_i.Rdata with 1 ≤ i ≤ 500 (folder Model_Predict/output)***

   Object: A list containing the elements: "*Mat*" and "*susp*".

Class: list

Object: Mat
Class: array

Description: An array of dimension 8 * 4533 * 138, with the numbers of pigs in each compartment (*S*, *E*, *Isc*, *Ic*, *R*), as well as the contribution of residues from dead pigs to transmission (*D*) and the numbers of new clinical cases (*Morb*) and new ASFV-related deaths (*Mort*), in each herd (4533) and at each time step (138).

Object: susp
Class: matrix

Description: A matrix of dimension 4533 * 138, with one row for each herd and one column for each time step, which takes value TRUE for herd *i* as long as the conditions for suspicion are met, and FALSE otherwise.