

ASF CHALLENGE MODEL DATA DOCUMENTATION UK TEAM

Phase 2

A. Data files - wild boar model

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

1. *locations.RDS*

Class: data frame

Description: Data frame with relevant information on all boar expected to be located within the landscape by day 0.

Column name	Short description	Additional comments
<i>X</i>	X-coordinate of boar location	
<i>Y</i>	Y-coordinate of boar location	
<i>date.removed</i>	Observed removal date	
<i>date.removed2</i>	Removal date in model simulation	
<i>date.infected:</i>	For positive boar, day of infection	Value is 0 for negative boar
<i>date.death.recovery</i>	For positive boar, expected day of death.	Value is 0 for negative boar
<i>date.end.infectious:</i>	For positive boar, expected day of for end of infectiousness.	Value is 0 for negative boar
<i>patch</i>	Patch to which boar belongs	Patches have IDs 1 through to 2500.
<i>infectionStatus</i>	Infection status of boar.	1 if infected and 0 if not infected
<i>index</i>	Unique ID of boar	
<i>det</i>	Detection mode	For observed boar locations, one of “NT”, “PT”, “AS”, and “NAS”, as recorded in the testing data. For simulated boar locations, “NA”.

<i>status</i>	Infection status as recorded in the observed data	For observed boar locations, one of “Positive” or “Negative”. For simulated boar locations, “Unknown”.
---------------	---	---

2. *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.

3. *in.fence.matrix.RDS*

Class: matrix

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

4. *in.fence.buffer.matrix.RDS*

Class: matrix

Description	Comment
Indicator matrix for whether a patch is either located in the fence or buffer zone	1 = Within fence or buffer zone 0 = Outside fence or buffer zone

5. *patchcentres.RDS*

Class: data frame

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

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

6. *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
<i>X</i>	X coordinate of patch	1 = Within fence-buffer 0 = Outside fence-buffer
<i>Y</i>	Y-coordinate of patch	
<i>region</i>	Administrative region in which patch center is located	
<i>number.of.boar</i>	Number of boar within patch	
<i>id</i>	Unique patch ID	In the range [1, 2500] as there are 2500 patches

7. *removals.RDS*

Class: data frame

Description: Daily data on boar removals within the area considered, from day 1 to day 80 (See Model description document for details on area considered.)

Column name	Short description	Additional comments
<i>in.zone</i>	Daily number of boar removed within zone by all methods	Removal methods considered are hunting, active search (AS) and passive surveillance (PS).
<i>in.zone.found</i>	Daily number of boar found by AS or PS within zone	
<i>in.zone.hunted</i>	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.
<i>out.zone</i>	Daily number of boar removed outside zone by all methods	
<i>out.zone.found</i>	Daily number of boar found by AS or PS outside zone	
<i>out.zone.hunted</i>	Daily number of boar hunted outside zone	
<i>total</i>	Total number of boar removed by all methods	

8. *nBoarsMatrix.RDS*

Class: matrix

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

9. *observed.positive.RDS*

Class: data frame

Description: Data on daily number of detected positive cases, from day 1 to day 80. This is only used directly in the parameter estimation process for computing summary statistics of the observed data.

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

10. *observed.positive.locations.RDS*

Class: list

Description: List containing coordinates of detected positive boar *by* day X , where $X = \{1, 2, \dots, 80\}$. The list has 80 elements (the first element corresponds to day 1 and the last element corresponds to day 80). 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	

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

A list (*sample_model_output_phase2.RDS) containing the elements:**

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

Column	Short description	Additional comments
<i>day</i>	Days for which simulation was performed	For parameter estimation, day ranged from 0 to 80.
<i>new.cases.wb</i>	Number of detected infected boar on day as determined by the model	
<i>cumulative.cases.wb</i>	Cumulative number of detected infected boar by 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.

locations: As described in section A.1. 1.

*Note that *sample_model_output_phase2.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, *param_estims_phase2.RDS*, is constituted of the following elements.

estims: Matrix with dimensions $N \times 1$, 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 threshold values). In Phase 2, $N = 500$.

mat: List with length equal to N ; each list element corresponds to an accepted particle. Each list element is a list of 80 indicator matrices, with the matrix in the first position corresponding to day 1 and the matrix in the last position corresponding to day 80. 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.

series: List with length equal to N ; each list element corresponds to an accepted particle. Each list element is a data frame of daily and cumulative counts of detected positive boar, as predicted by the accepted simulation. The columns of the data frame are described below.

Column	Short description	Additional comments
<i>day</i>	Days for which simulation was performed; day 81 to day 110	
<i>new.cases.wb</i>	Number of detected infected boar on day as determined by the model	

<i>cumulative.cases.wb</i>	Cumulative number of detected infected boar by day as determined by the model	
----------------------------	---	--

locations: List with length equal to N ; each list element corresponds to an accepted particle. Each list element is a data frame of boar locations (those in the landscape by day 0) as well as infection status and transition dates (if infected), as predicted by the accepted simulation. The columns of this data frame are as described in Section A.1. 1.

- The list, *param_estims_phase2_estims_only.RDS* holds the object *estims*, explained above.

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 110 under the increased hunting pressure and scenario and normal hunting pressure scenarios

The data files containing the predictions are:

File	Description	Details
<i>forward.run.increased.pressure.110.RDS</i>	Predictions from day 81 to day 110 under the increased hunting pressure scenario	<p>A list with the elements:</p> <p>time.series: List with length N, with each element being the prediction made by one accepted particle. Each list element is a data frame of daily and cumulative counts of detected positive boar over 30 days, from day 81 to day 110. Column descriptions are as described in <i>series</i> in Section A.3.</p> <p>status.matrices: List with length N, with each element being the prediction made by one accepted particle. Each list element is a list of 30 indicator matrices, with the matrix in the first position corresponding to day 81 and the matrix in the last position corresponding to day 110.</p>

		locations: List with length N . Each list element is a data frame of boar locations, infection status and transition (if infected) by day 110, as predicted by the accepted simulation.
<i>forward.run.normal.pressure.110.RDS</i>	Predictions from day 60 to day 110 under the normal hunting pressure scenario	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 81 up to day 110.

Column name	Short description	Additional comments
<i>source</i>	Unique identification number of the pig site that moved the pigs	Same as in moves_Players_day_80.csv
<i>dest</i>	Unique identification number of the pig site that received the pigs	
<i>source.type</i>	Production type of source herd	
<i>dest.type</i>	Production type of dest herd	
<i>date</i>	Day of the shipment	
<i>source.size</i>	Number of pigs in the source herd	Same as in herds_day_80.csv
<i>dest.size</i>	Number of pigs in the dest herd	Same as in herds_day_80.csv
<i>source.com</i>	Activity of the source herd (commercial/backyard)	1 for commercial sites; 0 for backyard sites
<i>dest.com</i>	Activity of the dest herd (commercial/backyard)	1 for commercial sites; 0 for backyard sites
<i>source.out</i>	Whether the pigs have access to an outdoor area or not in the source herd	1 for access to outdoor; 0 otherwise
<i>dest.out</i>	Whether the pigs have access to an outdoor area or not in the dest herd	1 for access to outdoor; 0 otherwise
<i>source.multi</i>	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
<i>dest.multi</i>	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
<i>multi</i>	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
<i>qty</i>	Total number of pigs of the shipment	Same as in moves_Players_day_80.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 $1 \leq i \leq 500$ (folder *Moves_Simul/sim*): see section B.2. Only columns "date", "source", "source.type", "dest", "dest.type" and "qty" (same as in [moves_Players_day_80.csv](#)) are used.

Folder *Model_Predict/input*:

- *patchcentres.RDS*: see section A.1.
- *param_estims_phase2.RDS*: see section A.3.
- *forward.run.increased.pressure.110.RDS*: see section A.4.
- *forward.run.normal.pressure.110.RDS*: see section A.4.

6. *forward.pressure.increasedhunting.110.RDS* (folder *Moves_Predict/input*)

Object: final.pressure

Class: list

Description: List of length 500, with each component containing a matrix of dimensions of $4537 * 110$ with the infectious pressures exerted by wild boar on each herd (4537) and at each time step in the wild boar model (110) for the scenario with increased hunting pressure.

7. *forward.pressure.normalhunting.110.RDS* (folder *Moves_Predict/input*)

Object: final.pressure

Class: list

Description: List of length 500, with each component containing a matrix of dimensions of $4537 * 110$ with the infectious pressures exerted by wild boar on each herd (4537) and at each time step in the wild boar model (110) for the scenario with normal hunting pressure.

B.4. Pig herd model output

8. *simul_i.Rdata* with $1 \leq i \leq 500$ (folder *Model_Predict/output*)

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

Class: list

Object: *Mat*

Class: array

Description: An array of dimension $8 * 4537 * 170$, 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 (4537) and at each time step (170).

Object: *susp*

Class: matrix

Description: A matrix of dimension $4537 * 170$, 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.