ASF CHALLENGE
MODEL DATA DOCUMENTATION
UK TEAM

Phase 3

A. Data files - wild boar model

A.1. Model inputs

1. *locations60.RDS*

   Class: data frame

Description: Data frame with relevant information on all boar expected to be located within the landscape by day 60.

Column name	Short description	Additional comments
X	X-coordinate of boar location	
Y	Y-coordinate of boar location	
date.removed	Observed removal date	
date.removed2	Removal date in model simulation	
date.infected:	For positive boar, day of infection	Value is 0 for negative boar
date.death.recovery	For positive boar, expected day of death.	Value is 0 for negative boar
date.end.infectious:	For positive boar, expected day of for end of infectiousness.	Value is 0 for negative boar
patch	Patch to which boar belongs	Patches have IDs 1 through to 2500.
infectionStatus	Infection status of boar.	1 if infected and 0 if not infected
index	Unique ID of boar	
carcass	Is boar a carcass?	1 if ever carcass and 0 if alive

2. *infectionStatusMatrix60.RDS*

Class: matrix

Description	Comment
Indicator matrix of infection statuses of patches by day 60	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
X	Patch coordinates in the X-direction	1 = Within fence-buffer
0 = Outside fence-buffer

Y	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
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

7. *removals.RDS*

Class: data frame

Description: Daily data on boar removals within the area considered, from day 1 to day 110 (See Model overview for details on area considered.)

Column name	Short description	Additional comments
in.zone.removed	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.removed	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.removed	Total number of boar removed by all methods	
in.zone.remaining	Number of boar remaining within the zone after the day's removals	The total number of boar in the island by day 0 was estimated using the 2019 hunting bag data and the estimate for the percentage of the wild boar population that is hunted in a typical season.
out.zone.remaining	Number of boar remaining outside the zone after the day's removals	
total.remaining	Total number of boar remaining in the zone after the day's removals	
total.hunted	Total number of hunted boar	
fraction.removed.in.zone	Fraction of boar removed within the zone by all methods	
fraction.removed.out.zone	Fraction of boar removed outside the zone by all methods	
in.zone.hunted.positive	Number of positive boar hunted within the zone	
in.zone.hunted.negative	Number of negative boar hunted within the zone	

out.zone.hunted.positive	Number of positive boar hunted outside the zone	
out.zone.hunted.negative	Number of negative boar hunted outside the zone	
fraction.found.in.zone	Fraction of boar found within the zone	
fraction.hunted.in.zone	Fraction of boar hunted within the zone	
fraction.hunted.out.zone	Fraction of boar hunted outside the zone	
fraction.found.out.zone	Fraction of boar found outside the zone	

8. *dist.patches.RDS*

Class: "dist"

Description: Distance matrix between all patch centres.

9. *nBoarsMatrix.RDS*

Class: matrix

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

10. *observed.positive.RDS*

Class: data frame

Description: Data on daily number of detected positive cases, from day 1 to day 110. 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

11. *observed.positive.locations.RDS*

Class: list

Description: List containing coordinates of detected positive boar *by* day *X*, where $X = \{1, 2, …, 110\}$. The list has 51 elements (the first corresponding to day 60 and the last element corresponding to day 110). 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

A list (*sample_model_ouptut_phase3.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 60 to 110. For predictions, day ranged from 111 to 230.
new.cases.wb	Number of detected infected boar on day as determined by the model	
cumulative.cases.wb	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_ouptut_phase3.RDS* is only intended to serve as a sample which reflects the <u>structure</u> of the output of the model. The output <u>values</u> should therefore not be regarded as typical model outcomes.

A.3. Parameter estimates

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

The list, *param_estims_phase3.RDS,* **is constituted of the following elements.**

estims: Matrix with dimensions N x 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 3, $N = 500$.

mat: List with length equal to N; each list element corresponds to an accepted particle. Each list element is a list of 51 indicator matrices, with the matrix in the first position corresponding to day 60 and the matrix in the last position corresponding to day 110. 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
day	Days for which simulation was performed; day 60 to day 110	
new.cases.wb	Number of detected infected boar on day as determined by the model	
cumulative.cases.wb	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 60) 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.

7

A.4. 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 230 under the increased hunting pressure scenario, and up to day 140 under the normal hunting pressure scenario.

The data files containing the predictions are:

File	Description	Details
forward.run.increased.pressure.140.RDS	Predictions from day 111 to day 140 under the increased hunting pressure scenario	A list with the elements: ***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 111 to day 140. Column descriptions are as described in ***series*** in Section A.3. ***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 111 and the matrix in the last position corresponding to day 140. ***locations:*** List with length *N.* Each list element is a data frame of boar locations, infection status and transition (if infected) by day 140, as predicted by the accepted simulation.
forward.run.normal.pressure.140.RDS	Predictions from day 60 to day 140 under the normal hunting pressure scenario	As above
forward.run.increased.pressure.230.RDS	Predictions from day 141 to day 230 under the	A list with the elements:

	increased hunting pressure scenario	**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 90 days: from day 141 to day 230. Column descriptions are as described in **series** in Section A.3.
		status.matrices: List with length *N,* with each element being the prediction made by one accepted particle. Each list element is a list of 90 indicator matrices, with the matrix in the first position corresponding to day 141 and the matrix in the last position corresponding to day 230.
		locations: List with length *N.* Each list element is a data frame of boar locations, infection status and transition dates (if infected) by day 230, as predicted by the accepted simulation.

A.5. Sensitivity analysis outputs

An output list as described in Section in A.2. but containing the following additional elements:

mir.value: the maximum infection range used in the simulation

alpha.value: the scale of the dispersal kernel (α) used in the simulation

infc.value: the duration of infectiousness in wild boar carcasses used in the simulation

infd.value: the duration of infectiousness in live boar used in the simulation

B. <u>Data files – pig herd model</u>

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 ≤ i ≤ 500 (folder Moves_Simul/sim)*

Object: mov.pred2
Class: data.frame

Description: Simulated pig movements from day 111 up to day 230.

Column name	Short description	Additional comments
source	Unique identification number of the pig site that moved the pigs	Same as in *moves_Players_day_110.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_110.csv*
dest.size	Number of pigs in the dest herd	Same as in *herds_day_110.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_110.csv*

4. *mov_Se_i.Rdata with 1 ≤ i ≤ 500 (folder Moves_Simul/Se)*

   Object: Se
   Class: numeric

Description: A vector containing the sensitivities of the simulated movement networks (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 networks). As we predict for four months (120 days), there are four values, one value for each simulated month (30 days) of movements.

5. *mov_Sp_i.Rdata with 1 ≤ i ≤ 500 (folder Moves_Simul/Sp)*

   Object: Sp
   Class: numeric

Description: A vector containing the specificities of the simulated movement networks (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 networks). As we predict for four months (120 days), there are four values, one value for each simulated month (30 days) of movements.

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_Players_day_110.csv*) are used.

Folder Model_Predict/input or Model_Predict_230/input:
  - *locations60.RDS:* see section A.1.
  - *patchcentres.RDS:* see section A.1.
  - *locations.day.59.RDS:* infection and location data on wild boar removed up to day 59. Same columns as locations60.RDS.
  - *param_estims.RDS:* see section A.3.
  - *forward.run.increased.pressure.140.RDS:* see section A.4.
  - *forward.run.normal.pressure.140.RDS:* see section A.4.
  - *forward.run.increased.pressure.230.RDS:* see section A.4.

B.4. Pig herd model output

6. ***simul_i.Rdata with 1 ≤ i ≤ 500 (folder Model_Predict/output or Model_Predict_230/output)***

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

Object: Mat
Class: array

Description: An array of dimension 8 * 4540 * 290, 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 (4540) and at each time step (290).

Object: susp
Class: matrix

Description: A matrix of dimension 4540 * 290, with one row for each herd and one column for each time step, which takes value TRUE for herd i suspected at time step t, and FALSE otherwise.

Object: CO
Class: matrix

Description: A matrix of dimension 4540 * 290, with one row for each herd and one column for each time step, which takes value TRUE for herd i confirmed infected at time step t, and FALSE otherwise.