

Prototype Biodiversity Digital Twin: Disease Outbreaks

Kate Ingenloff[‡], Syrine Ben Aziza[§], Claus Weiland[¶]#, Nikoletta Nikolova[§], Hans-Hermann Thulke[¶], Martin Lange[¶], Adam Reichold[¶], Dmitry Schigel[‡]

[‡] Global Biodiversity Information Facility - Secretariat, Copenhagen Ø, Denmark

[§] TNO, The Hague, Netherlands

[|] Senckenberg – Leibniz Institution for Biodiversity and Earth System Research, Frankfurt am Main, Germany

[¶] Senckenberg Research Institute and Natural History Museum, Frankfurt am Main, Germany

[#] DiSSCo-D, Frankfurt am Main, Germany

[□] Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

Corresponding author: Kate Ingenloff (kathryn.ingenloff@gmail.com)

Academic editor: Christoph Wohner

Abstract

African swine fever is a transmissible virus impacting wild and domestic swine populations. In Europe, it is non-native and the recently introduced genotype affects wild boar populations with occasional outbreaks in domestic pigs. The ability to predict short-term spatial dynamics of this disease will greatly improve our ability to control and limit future spread of the virus. The BioDT disease outbreaks prototype, currently in development, implements an individual-based landscape-level mechanistic model as a digital twin aimed at providing informed support for management decisions in response to the spread of African swine fever in European wild boar populations.

Keywords

wild boar, *Sus scrofa*, African swine fever, ecological modelling, outbreak management, mechanistic model

Introduction

African swine fever (ASF) is a transmittable lethal infection of wild boar and domestic swine (*Sus scrofa*) caused by a DNA virus of the genus *Asfivirus* (family *Asfarviridae*; Gabriel et al. 2011). Research suggests that, in Europe, traditional transmission routes involving blood-sucking soft ticks of the genus *Ornithodoros* do not play a role in the ASF infection cycle (Pietschmann et al. 2016). Instead, animals are infected through direct contact with other infected animals or contaminated carcasses, food, or equipment (Gaudreault et al. 2020). Although it is harmless to people, the virus can have a significant

economic impact. There is no vaccine or cure despite ongoing research (Cwynar et al. 2019, Gaudreault et al. 2020) limiting control measures to early detection of infection, physical separation of healthy from diseased individuals, and culling of infected animals (European Food Safety Authority 2014, Jurado et al. 2018).

First identified in Kenya in 1921 (Montgomery 1921), multiple genotypes of the ASF virus are found in 35 African countries (Gaudreault et al. 2020) and one form, Genotype 2, affects over 60 countries across the rest of the world (Gaudreault et al. 2020) as a result of human-facilitated continental translocation in 2007. In Europe, the virus is restricted mostly to wild boar populations with isolated infections occurring in domestic pigs (Cwynar et al. 2019). The continued occurrence of the disease in wild boar populations and the limited number of available control measures represents a particular challenge for agriculture at a global level, in particular the pig farming industry (Gallardo et al. 2015). As such, rapid identification of the best control measure(s) during an emerging outbreak is critical. Mechanistic models are valuable tools in estimating the trajectory of an ongoing disease outbreak, assessing efficacy of control measures, and predicting potential future outbreaks (Thulke 2011, Lange et al. 2018, Hayes et al. 2021).

The BioDT disease outbreaks prototype digital twin (pDT) aims to establish a predictive modelling tool that can be used in acute ASF outbreaks by decision-making bodies (<https://biodt.eu/use-cases/disease-outbreaks>). It implements an established individual-based, spatially-explicit mechanistic model at the landscape scale to simulate the spread and potential control of ASF in European wild boar populations (Lange and Thulke 2017). The general model framework (<https://ecoepi.eu/ASFWB/>) is already used to guide decision-making in a prospective manner (e.g. European Food Safety Authority et al. 2017, Halasa et al. 2019, European Food Safety Authority et al. 2021), but people or entities interested in accessing the model must contact the creators directly. Implementation of the model as a digital twin via the BioDT project will make the model available as needed to public health officials supporting responsive control efforts during ASF outbreaks. Here, we provide a snapshot of progress in adapting the model to a digital twin, highlighting the design and potential impact of the digital twin as well as current challenges.

Objectives

The purpose of this pDT is to provide information for data-driven responses in managing the spread of African swine fever in European wild boar populations. With each reported detection of ASF, the model is updated and rerun to produce more robust predictions of infection risk and the effectiveness of spatially-explicit control measures permitting routine update of the most suitable virus containment decisions during an ASF outbreak. Although there are multiple potential stakeholders and user groups, the initial implementation of this BioDT prototype digital twin focuses on public health decision-makers and people involved in decision support.

Data

The wild boar-African swine fever model requires two types of data: static and dynamic (Table 1).

Static data consist of habitat or landscape structure data. These data describe the landscape on which the model simulations take place. Provision of a habitat structure data layer triggers the first run of a wild boar–ASF model simulation. Currently, the static data layer must be created by the user or obtained from ENetWild (ENETWILD-Consortium et al. 2020) and manually supplied to the model by the user.

Dynamic data can be supplied to the model any time new information becomes available, triggering the model to rerun and return updated results based on the new information. These data include wild boar geolocation and ASF infection information as well as virus control scenarios. Geolocation data include the location (as GPS coordinates), time and date a boar (or group of wild boar) was identified. These data may also supply the model with an animal's infection status (is the animal healthy, infected, or recovered?), as well as other demographic information about an animal, such as its age, sex, and vitality (whether or not the animal is alive or dead). Proposed virus control measures or treatment scenarios are location-specific measures designed to halt or contain the infection including barriers designed to separate infected from non-infected individuals, local depopulation efforts and carcass retrieval. A user will be able to supply the model with multiple proposed control measures to identify the best control scenario. In an ideal scenario, wild boar geolocation data and ASF infection reports would be directly updated from partnered research institutes (RI), such as the Global Biodiversity Information Facility (GBIF); however, presently, all data must be supplied to the model by the user.

Model

Fig. 1 provides an overview of the wild boar-ASF pDT . The model is stochastic, spatially- and temporally-explicit, and individually-based. It incorporates wild boar ecology and ASF epidemiology to simulate infection dynamics in the wild boar population. The first run of a wild boar-ASF model simulation is triggered when habitat structure data are supplied to the model. Each time additional information is provided (see dynamic data; Table 1), the model reruns, incorporating the new data and returning an updated, dynamic prediction of African swine fever infection including identification of active and recovered cases as well as incidents of wild boar death due to the virus (Lange and Thulke 2017). The project intends to provide users with a standardised suite of model output data, although the composition of this downloadable data packet has yet to be agreed upon and implemented in the code.

The model was designed using a modular format (known as a modularised entity component system or ECS). This structure ensures that all model sub-modules have the

same structure and interface. The core code of each module is written in the Rust programming language (<https://www.rust-lang.org/>); and each module has a 'wrapper' written in Python (<https://www.python.org/>). This setup allows for a fully customisable model module configuration that can be adapted to suit each modeller's needs. With the assistance of H.-H. Thulke and M. Lange, the pDT team selected a default configuration for use in the initial pDT implementation.

At the close of the second year of the BioDT project, development of the wild boar-ASF pDT is well underway, but not yet complete (see Table 2 for a full list of pDT implementation tasks and their current status). Initial testing and exploration of the model in the local environment is complete and the process of migrating the tested minimum working model to the LUMI supercomputing environment (<https://lumi-supercomputer.eu/>) initiated. Full pDT functionality, however, requires two additional model development efforts: (1) automation of proposed infection control scenarios ingestion and manipulation and (2) adaptation of the code to return the standardised suite of model outputs. Resource availability to implement these model improvements is presently unclear.

FAIRness

Ensuring that project data and model outputs are FAIR (Findable, Accessible, Interoperable and Reusable) is a high priority for the BioDT initiative as FAIR principles are critical for enabling autonomous or semi-autonomous navigation and processing of data in large integrated data spaces by machines (Wilkinson et al. 2016, Duatis Juarez et al. 2023). To enable this functionality, project data (input and output) will need to be upgraded to include explicit semantics describing how the data can be automatically processed and integrated into an ecosystem of FAIR-enabling services which implement the required policies, rules, procedures and infrastructures for the operation concerned (European Commission and Directorate-General for Research and Innovation 2018). The wild boar-ASF team is working with BioDT FAIR data specialists to incorporate this FAIR Digital Object (FDO, Wittenburg et al. 2023) concept, although the implementation workflow has not yet been established. Following this, model data, outputs, and metadata will be formatted as lightweight FDOs to facilitate sharing for re-use using [WorkflowHub](#) (Soiland-Reyes 2023).

Performance

The wild boar-ASF model is not computationally intensive. The initial model implementation runs easily on a processor Intel(R) Core(TM) i7-8665U CPU @ 1.90GHz, 2112 Mhz, 4 Core with 8 Logical Processors running Windows version 10.0.19045 Build 19045. The simulation's current computational profile features an overall execution time of 8.07 seconds and a granularity of 15.51 milliseconds per simulation tick over a span of 520 ticks. The spatial range is bounded by [4_506_779, 3_052_929, 4_855_174, 3_353_689] with a resolution of 2000 m. Initial agents are defined with a release factor of

5.0 and the mutation probability is set to $1e-2$. These metrics suggest that the current model instance is less computationally demanding than what is expected in large-scale deployments, which typically involve higher resolutions and more extensive agent-based modelling. However, when the model is applied to continental-scale problems with larger datasets, higher agent counts and finer spatial resolutions, the existing architecture may limit performance efficiency of large multi-repetition experiments.

Introducing parallelisation when the model is migrated to the LUMI supercomputing environment will be essential as a scalability measure, as well as a means of enhancing computational throughput. Should a user wish to explore multiple virus barrier scenarios at the same time, parallel processing will enable multiple, independent model simulations to run simultaneously, reducing overall runtime and making efficient use of computational resources. This will be particularly beneficial during acute ASF outbreaks when multiple intervention strategies must be rapidly evaluated.

Interface and outputs

A user interface (UI) for the pDT is currently in development with project collaborators at CSC - IT Center for Science LT (<https://www.csc.fi>) and VSB - Technical University of Ostrava (<https://www.vsb.cz>). Individual users will have their own password-protected account and the UI will facilitate three user capabilities (see Fig. 1). Firstly, users will have access to a webform by which they can upload data to initiate or supplement a simulation. The UI will also ensure that users can access their archived, version-controlled input data and model outputs which will be stored in a user-specific, project-facilitated repository. Additionally, users will be able to view, archive, and download a standardised suite of model output information. While the details have yet to be finalised, this information should include summaries of projected outcomes, static model graphics (e.g., maps, graphs) summarising projected outcomes, and spatially dynamic views of model predictions through time. Dynamic outputs will be downloadable in movie, gif, or similar format; static data will include a PDF summary output and individual, high-resolution graphics (e.g., maps, graphs).

Integration and sustainability

There are no plans at the moment for any connection with the Destination Earth (DestinE, <https://digital-strategy.ec.europa.eu/en/policies/destination-earth>) or the European Open Science Cloud (EOSC, <https://eosc.eu/eosc-about/>), nor is the pDT supported by any of the project biodiversity research institutes (RI). Additionally, the sustainability of the project remains in question. With only one year of support left under the broader BioDT project, the wild boar-ASF digital twin will require funding and support to keep it running. Support will be necessary to cover the technical infrastructure of the project and provide a means of funding on-going model development and implementing improvements to the user interface.

Application and impact

The wild boar-ASF prototype digital twin aims to address an ongoing need in tracking and responding to wildlife disease outbreaks, such as African swine fever. The ability to predict short-term spatial dynamics of ASF will improve efforts to control and limit potential spread. However, the initial implementation of this digital twin only simulates ASF within the wild boar population and some aspects of model development remain incomplete. The major barriers to progress are a lack of adequate funding streams to support long-term pDT infrastructure and model development and a lack of openly available disease host and infection data. Despite this, the project still hopes to eventually expand the model to simulate and assess the risk of ASF spillover from wild boar to domestic pig populations.

Spillover of ASF infection into domestic pig populations has significant economic and food-security impacts. The ability to model this potential on-demand and in real-time would minimise risk of transmission and protect the livelihoods of ranchers supplying pork and other pig products. In addition to requiring on-going support for model development and maintenance as information about wild boar-domestic swine-ASF improves, implementation of a full spillover risk model as a digital twin will present a new suite of challenges including significantly increased computational demand (the model linking the wild boar-ASF and domestic swine-ASF models would need to simulate feedback interactions) and a need for much greater access to near real-time wild boar, ASF infection, and domestic pig data via project RIs.

While the wild boar-ASF model is fully documented and well-established as a model-based policy support tool, the process of adapting it to a digital twin highlighted the need for greater generalisation of the model to effectively adapt it into an independent forecasting tool. Unfortunately, some of these model development needs (specifically automation of barrier data ingestion and standardisation of the model output data packet) are currently on hold as project resources for the modelling team are capped for the year. The pDT team hopes to resolve this before long so that model development can continue. In the meantime, development of the user interface and testing and implementation of the model in LUMI will continue. The team will also continue to address the dearth of available wild boar geolocation and ASF infection data in project RIs.

Acknowledgements

This study has received funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement No 101057437 (BioDT project, <https://doi.org/10.3030/101057437>). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Cwynar P, Stojkov J, Wlazlak K (2019) African Swine Fever Status in Europe. *Viruses* 11 (4). <https://doi.org/10.3390/v11040310>
- Duatis Juarez J, Schick M, Puechmaille D, Stoicescu M, Saulyak B (2023) Destination Earth Data Lake. EGU General Assembly, Vienna, Austria, 24–28 Apr 2023. <https://doi.org/10.5194/egusphere-egu23-7177>
- ENETWILD-Consortium, Acevedo P, Croft S, Smith G, Blanco-Aguilar JA, Fernández-López J, Scandura M, Apollonio M, Ferroglio E, Keuling O, Sange M, Zanet S, Brivio F, Podgorski T, Petrovic K, Soriguer R, Vicente J (2020) Model outputs for update of occurrence and hunting yield-based data models for wild boar at European scale: new approach to handle the bioregion effect, May 2020 update. Zenodo <https://doi.org/10.5281/zenodo.3823925>
- European Commission, Directorate-General for Research and Innovation (2018) Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data. Publications Office. <https://doi.org/10.2777/1524>
- European Food Safety Authority (2014) Evaluation of possible mitigation measures to prevent introduction and spread of African swine fever virus through wild boar. *EFSA Journal* 12: 3616. <https://doi.org/10.2903/j.efsa.2014.3616>
- European Food Safety Authority, Depner K, Gortazar C, Guberti V, Masiulis M, More S, Ojševskis E, Thulke HH, Viltrop A, Woźniakowski G, Cortiñas Abrahantes J, Gogin A, Verdonck F, Dhollander S (2017) Scientific Report on the epidemiological analyses of African swine fever in the Baltic States and Poland. *EFSA Journal*. *EFSA Journal* 15: 5068. <https://doi.org/10.2903/j.efsa.2017.5068>
- European Food Safety Authority, Nielsen SS, Alvarez J, Bicout DJ, Calistri P, Depner K, Drewe JA, Garin-Bastuji B, Gonzales Rojas JL, Gortazar Schmidt C, Herskin M, Michel V, Miranda Chueca MÁ, Pasquali P, Roberts HC, Sihvonen LH, Spoolder H, Stahl K, Velarde A, Winckler C, Abrahantes JC, Dhollander S, Ivanciu C, Papanikolaou A, Van der Stede Y, Blome S, Guberti V, Loi F, More S, Olsevskis E, Thulke HH, Viltrop A (2021) ASF Exit Strategy: Providing cumulative evidence of the absence of African swine fever virus circulation in wild boar populations using standard surveillance measures. *EFSA Journal* 19: 6419. <https://doi.org/10.2903/j.efsa.2021.6419>
- Gabriel C, Blome S, Malogolovkin A, Parilov S, Kolbasov D, Teifke JP, Beer M (2011) Characterization of African swine fever virus Caucasus isolate in European wild boars. *Emerging Infectious Disease* 17: 2342-2345. <https://doi.org/10.3201/eid1712.110430>
- Gallardo MC, Reoyo Ad, Fernández-Pinero J, Iglesias I, Muñoz MJ, Arias ML (2015) African swine fever: a global view of the current challenge. *Porcine Health Management* 1: 21. <https://doi.org/10.1186/s40813-015-0013>

- Gaudreault N, Madden D, Wilson W, Trujillo J, Richt J (2020) African Swine Fever Virus: An Emerging DNA Arbovirus. *Frontiers in Veterinary Science* 7 <https://doi.org/10.3389/fvets.2020.00215>
- Halasa T, Boklund A, Bøtner A, Mortensen S, Kjær LJ (2019) Simulation of transmission and persistence of African swine fever in wild boar in Denmark. *Preventative Veterinary Medicine* 167: 68-79. <https://doi.org/10.1016/j.prevetmed.2019.03.028>
- Hayes B, Andraud M, Salazar L, Rose N, Vergne T (2021) Mechanistic modelling of African swine fever: A systematic review. *Preventive Veterinary Medicine* 191: 105358. <https://doi.org/10.1016/j.prevetmed.2021.105358>
- Jurado C, Martínez-Avilés M, De La Torre A, Štukelj M, de Carvalho Ferreira HC, Cerioli M, Sánchez-Vizcaino JM, Bellini S (2018) Relevant Measures to Prevent the Spread of African Swine Fever in the European Union Domestic Pig Sector. *Frontiers in Veterinary Science* 5: 77. <https://doi.org/10.3389/fvets.2018.00077>
- Lange M, Thulke HH (2017) Elucidating transmission parameters of African swine fever through wild boar carcasses by combining spatio-temporal notification data and agent-based modelling. *Stochastic Environmental Research and Risk Assessment* 31 (2): 379-391. <https://doi.org/10.1007/s00477-016-1358-8>
- Lange M, Guberti V, Thulke HH (2018) Understanding ASF spread and emergency control concepts in wild boar populations using individual-based modelling and spatio-temporal surveillance data. *EFSA Supporting Publications* 15 (11). <https://doi.org/10.2903/sp.efsa.2018.en-1521>
- Montgomery RE (1921) On A Form of Swine Fever Occurring in British East Africa (Kenya Colony). *Journal of Comparative Pathology and Therapeutics* 34: 159-191. [https://doi.org/10.1016/s0368-1742\(21\)80031-4](https://doi.org/10.1016/s0368-1742(21)80031-4)
- Pietschmann J, Mur L, Blome S, Beer M, Pérez-Sánchez R, Oleaga A, Sánchez-Vizcaino JM (2016) African swine fever virus transmission cycles in Central Europe: Evaluation of wild boar-soft tick contacts through detection of antibodies against *Ornithodoros erraticus* saliva antigen. *BMC Veterinary Research* 12 (1). <https://doi.org/10.1186/s12917-015-0629-9>
- Soiland-Reyes S (2023) FAIR Research Objects and Computational Workflows – A Linked Data Approach. Zenodo <https://doi.org/10.5281/zenodo.8113626>
- Thulke HH (2011) Application of recent approaches in modelling for Animal Health. *Preventive Veterinary Medicine* 99: 1-3. <https://doi.org/10.1016/j.prevetmed.2011.01.007>
- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, 't Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M, Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3 (1). <https://doi.org/10.1038/sdata.2016.18>
- Wittenburg P, Schwarzmann U, Bianchi C, Weiland C (2023) FDOs to Enable Cross-Silo Work. *Proceedings of the Conference on Research Data Infrastructure* 1 <https://doi.org/10.52825/cordi.v1i.263>

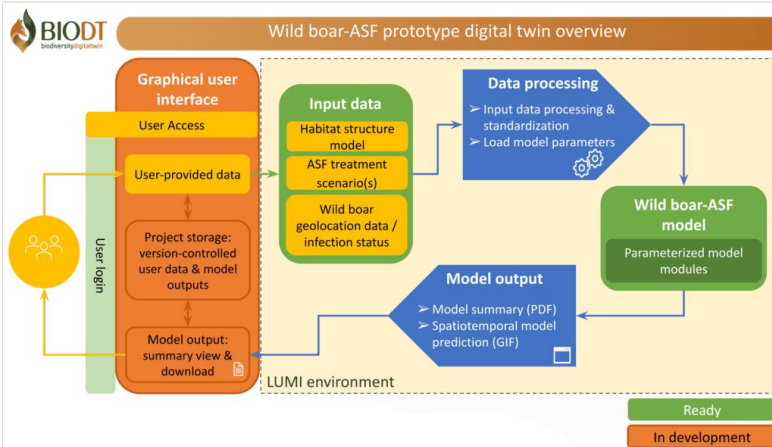


Figure 1. Schematic overview of the wild boar-ASF prototype digital twin with current status of major components.

Table 1.

Overview of static and dynamic data, known data sources and anticipated data formats.

Static/Dynamic	Data	Data source(s)	Data format(s)
Static	Habitat structure model	ENetWild , User-provided	Ascii; GeoTIFF; Raster
Dynamic	Wild boar geolocation information	ENetWild , User-provided	Text file; Comma-separated file
Dynamic	Wild boar infection status	User-provided	Text file; Comma-separated file
Dynamic	ASF treatment scenarios (barriers)	User-provided	Shapefile

Table 2.

Overview of the wild boar-ASF prototype digital twin development tasks and their current status.

Category	Task	Status
General	Obtain minimum working model module configuration and test data from project collaborators	Complete
Data	Fair data object creation	Incomplete
Model outputs	Select standardised suite of static and dynamic model outputs	Not started
Model – local environment	Exploration and testing of test data and minimum working model modular components in the local environment (model workflow)	Complete
	Model development: automation of treatment data ingestion/manipulation	Not started
	Model development: code adaptation to provide standardised model outputs	Not started
Model – LUMI	Model code made available to LUMI contacts	Complete
	Model deployment on LUMI supercomputing environment	Not started
	Model automation, parallelisation, and testing on LUMI	Not started
pDT user access	User Interface storyline drafting	Complete
	User Interface design and testing	Incomplete
	User-specific, version-controlled storage	Incomplete