A Digital Twin Simulator Approach as a Support to Develop an Integrated Observatory of the Epidemic Risk in a Rural Community in Senegal

Jean Le Fur¹, Moussa Sall² and Jean-Marie Dembele²

¹Institut de Recherche pour le Développement (IRD), Centre de Biologie pour la Gestion des Populations (CBGP), Campus Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez, France
²Dépt. Informatique, Univ. G. Berger/Saint-Louis, Senegal

Keywords: Digital Twin, Epidemic Risk, Agent-Based Model, Data Driven Approach, EcoHealth Approach, Synthetic Ecology, Complex System.

Abstract: Following the contemporary epidemiologic approach known as EcoHealth, the study of an epidemic risk must consider and integrate the whole set of actors, factors and environments bound to the transmission of infectious diseases. In this study, we propose using a mechanistically rich digital twin simulator as a tool to facilitate this integration with the addition of a functional and dynamic dimension. The selected case study is the monitoring of the risk associated with ticks and rodents in a rural community in the Sahelian region of Senegal. To construct the digital twin, we iteratively went back and forth between field data collection and computer transcription of knowledge. Thanks to the high resolution afforded by the digital twin approach, the simulator enables the study of city-scale activity patterns as well as interactions between ticks, rodents, cats, and humans that occur within habitation rooms and shops. In addition to (i) being able to provide dynamic integrated support for the collected multidisciplinary knowledge, the digital twin realism provides (ii) an appropriate medium for communicating results to non-expert populations and (iii) a useful tool for monitoring and adjusting the observatory’s data collection protocols. The model’s complexity presents calibration challenges that are discussed.

1 INTRODUCTION

Approximately two-thirds of known infectious diseases in humans are zoonoses, which are diseases whose pathogens are transmitted between vertebrate animals and humans (Jones et al., 2008). These diseases typically entail a multitude of interacting actors. In the typical instance of tick-borne diseases such as Lyme disease (Adrion, 2015) or Borreliosis (Elbir et al., 2015), the interacting agents are pathogenic bacteria, tick vectors of these diseases, intermediate rodent hosts that serve as reservoirs for pathogens, predators (e.g., cats) that regulate populations, and finally, humans susceptible to infection. All of these agents interact within diverse environments on varied length scales (e.g., nests or burrows of rodents, rooms or market places). The inseparable interconnection of the components within such complex systems is evident and requires integrated approaches to human and animal health and their respective social and environmental contexts (Zinsstag et al., 2011). The so-called EcoHealth (Listitza and Wolbring, 2018) or OneHealth (Mencke, 2013) approaches suggest jointly taking into account the knowledge produced by various thematic disciplines to better account for the processes at work in the transmission or lack thereof of a disease. However, a number of authors (e.g., Lefrançois et al., 2023, Rotureau et al., 2022) point out that it is still challenging to implement such multidisciplinary approaches due to the number of factors to be considered, scientific disciplines to be involved and the difficulty of integrating them. In this process of articulating acquired knowledge, modern modelling and simulation tools, and particularly the object paradigm derived from computer science, are important assets. Indeed, in terms of the study of such complex systems, numerical experiments allow for the articulation of...
various types of determinants and the comprehension of the potential outcomes of experiments that are difficult or impossible to conduct in situ (Auffray et al., 2011).

To implement such an approach, the observed reality must be represented in the most accurate manner. This is the objective of the approach in terms of digital twins (Tao et al., 2022). Usually devised to account for complicated (industrial processes, factories, robots) devices, the approach in terms of digital twin can also be considered as a metaphor that can be adapted to account for complex phenomena (e.g., Nie and al., 2023) including those involving actors in society (Grieves and Vickers, 2017 in White et al., 2021). The major outcome expected with the use of such an approach is to produce artefacts that are close enough to known reality to allow the processes to be studied in as much detail as possible, particularly at a very high spatial resolution. Since the occurrence or not of an infection of humans, rodents, ticks depends on the unique configuration of each dwelling, room, or activity, the digital twin paradigm was felt as a promising approach to deal with an EcoHealth approach of the epidemic risk.

In this work, we present the digital twin approach that we have developed to represent and exploit the knowledge acquired in the context of developing an EcoHealth-type integrated observatory of the epidemic risk in a rural African community.

2 MATERIAL AND METHOD

2.1 Model Purpose

We propose in this work to set up an integrative tool allowing to articulate abiotic, trophic, physiological, behavioural, social, demographic and environmental factors involved in the spread of epidemics in a typical rural Sahelian community in Senegal. However, rather than storing the collected data within a static database management system, the knowledge gathered on the various aspects of the system studied is placed into dynamic relation through a simulator that accounts for behaviours and interactions. The issue is concretely of developing a model as realistic as possible by using a so-called "mechanistically rich" approach (DeAngelis and Mooij, 2003) bringing together the knowledge from diverse thematic disciplines such as bio-ecology, parasitology, geography urban, social and human sciences.

The approach is also founded on a data-driven strategy for which the model's primary purpose is not to replicate the actual dynamics, but rather to serve as a repository for retrieving domain-specific knowledge. It aims, for instance, to identify the components for which there is a knowledge gap and to serve as a forum for discussion between scientists from different disciplines as well as between scientists and stakeholders (population, decision-makers) to ensure that the observatory and its protocols are continuously improved. The subsequent objective is to gain a simulation tool capable of simulating a variety of epidemic risk evolution scenarios (from pathogens to humans, from shops to burrows).

2.2 Case Study and Modelling Platform

The pilot site that has been selected to develop the EcoHealth observatory is the town of Dodel (16°29'10.1"N 14°25'56.5"W), a typical rural community of the Sahel in northern Senegal. In the chosen case study, the emphasis was placed on the risk associated with the transmission of Borreliosis or relapsing tick-borne fever, as described in the introduction.

The simulation model used as support for this work has already been developed and presented in Le Fur et al. (2017, 2021). The model is developed using Java and the agent-based technology. It is developed using the Repast-Simphony platform (North et al., 2013). Based on a bio-inspired approach (Le Fur et al., 2023), it enables the representation of synthetic ecologies that include a variety of animal and human biological agents with their own characteristics, capacities for observation, deliberation, action and interaction. The model is based on a parsimonious data system using chronograms of various events. It uses a double formalisation of space with a continuous Euclidean space for processing moves combined with a discrete grid for reifying the environment's constituent objects. The time step of the model is configurable in the sense that changes in time step are passed on to all the formalized dynamic processes.

2.3 Achievement

The development of a digital twin of the city of Dodel has been predicated on the principle of continuous improvement. It entails to incrementally build the model by performing round trips between the acquisition of multidisciplinary data in the field and the digital transcription of the collected information. The simulator model hence progressively gathers and connects, in an integrated and dynamic manner, the most possible complete set of available data and
knowledge on the various factors influencing the risk of epidemics. In this section, for each of the acquisition stages, we detail the fieldwork that was performed, then describe the formalisation that was retained and developed in order to construct, in parallel, the multidisciplinary observatory and its supporting simulator.

### 2.3.1 Cartography

A high-resolution map of the study area was the initial step. A team consisting of a modeller and geographer conducted a two-week survey of the entire city centre. Concessions, buildings, rooms and remarkable points have been mapped. A survey was conducted on each property to determine the number of structures, the type of fence, and, most importantly, the number, type, and purpose of each room. In accordance with the logic of the digital twin, the readings included the interior of homes in order to have the most accurate representation of the rooms and corridors where humans, pathogens, reservoirs, or vectors may come into contact.

The collected data were then georeferenced with a geographic information system, rasterized with a 1 m resolution, and incorporated into the simulator as a grid where each cell represents an object with attributes (type of place, presence of food, etc.). A dedicated algorithm then identified contiguous clusters of cells in order to determine the functional level of the various spaces (ex: all cells contained in a room are gathered in a higher level object of room type).

### 2.3.2 Population Census

The resident population census was conducted at the same time. This effort resulted in the identification of all downtown residents and their respective occupations. Each resident was then reified in the simulator as a human type agent positioned in her/his residence (or her/his shop for merchants).

Following the bio-inspired approach described in Le Fur et al. (2023), each inhabitant is characterised by attributes unique to either agents (name, age, location, energy), containers (containing container, containers contained, including pathogens), animals (speed, sensing, trapped, etc.), or mammals (mating latency, pregnancy, sexual maturity, suckling child, gestation length, etc.). The value of each attribute is provided either from data collected in the field (sex, age, pregnancy status, etc.) or from literature (max age, speed, sensing, etc.).

#### 2.3.3 Rodent Trapping

Secondly, a new ten days survey was carried out in the city to sample the main animal actors involved in the epidemic risk. Traps were placed in selected rooms several nights in a row and the rodents present were captured. Following a standardised protocol, each mouse capture led to biometric measurements, observation of the genitals and an autopsy with samples taken from different organs. In the model each rodent observation was implemented as it has been for humans by characterizing each agent by its class attributes, acquired in the field (sex, estimated age, pregnancy status in the case of females,...) or recovered from the literature.

The behaviour of house mice, like that of other mobile organisms in the model except humans (see below) is driven by a desire - perception - deliberation - decision - action type system (Le Fur et al., 2023). Concerning mice, modelling has also been the subject of specific treatment to account for their particular behaviour which compel them to move along walls during their foraging activity. An algorithm has been developed specifically for this purpose (Sall et al., 2019).

#### 2.3.4 Ticks

The protocol for collecting ticks involved locating rodent burrows and crevices inside homes and in the courtyards of compounds visited for rodent sampling, and then vacuuming the contents of these holes with a thermal vacuum cleaner. Each sample was then visually inspected for ticks, and the collected ticks were then placed in tubes.

The identified ticks were reified in the model using an object class inheriting from the animal category (sensing, velocity, etc.) with tick-specific properties. In this case, additional attributes have been added to represent the different stasis through which these animals pass (egg, larvae, nymph, and adult) as well as the processes of deliberation - decision (waiting in a burrow, climbing on a rodent or a human) and the timings allowing them to switch from one to the other by incorporating three specific attributes, namely the duration of a meal, the duration of hibernation after a meal, and the number of meals (McCoy and Boulanger, 2015).

#### 2.3.5 Cats

During the survey to trap rodents and sample ticks, the presence of cats was estimated based on direct observations coupled with a questionnaire administered to inhabitants. As precedently, the
identified cats were reified into a distinct class inheriting the traits of mobile mammals completed with a relatively simple specific behaviour (sleep for 14 hours during the day, one meal at home per day and a search activity for rodents in the streets of the village at night).

2.3.6 Bacteria

Finally, the rodent and tick samples were sent to a laboratory for bacterial strain identification (Ouarti et al., 2022). When infected rodents or ticks were identified, bacteria-like agents were reified and introduced into the mouse or tick agents, respectively, within the simulator. There are no agent-specific attribute or behaviour for bacteria which are passively transmitted from inside one agent to inside another during a tick bite in the current model (mouse to tick, tick to mouse, tick to human).

2.3.7 Human Activity

After completing this initial inventory of the present protagonists, a decisive new field survey was conducted to comprehend the daily rhythm of the city. The purpose then was to complete the model in order to determine and quantify the human activity periods that are susceptible to epidemic risk (house mouse movements, tick bites). Over the course of eight days, every house in the city centre was re-visited, and the residents (adults and adolescents) were questioned individually about their daily activities, hour by hour, as well as the time and place where they slept. These surveys resulted in the recording of 3,519 instances of activity involving 489 inhabitants and hence enabling a nearly exhaustive description of daily human activity in the city center. The wide variety of activity obtained was synthesized following a logic of appreciation of the possible interaction between humans on the one hand and rodents or ticks that can transmit diseases on the other. Each of these activities has thus been categorised into four main types: (i) movement corresponding to daily activity including travel, (ii) wakefulness corresponding to the times when the inhabitants remain steady but with their eyes open and can perceive their environment (prayer, internet, etc.); (iii) rest which is associated with sleep (night, siesta) and during which the inhabitants do not perceive their environment, finally (iv) meals during which food can constitute an olfactory stimulation and behavioural changes for rodents. The 3,519 activity instances were thus compiled and transcribed in this way before being added to the simulator's database.

Given the explicit nature of the activity sequences obtained as a result of this work, the human agents in the model were not subjected to a process of deliberation for their actions, as was the case for mice, cats, and ticks; instead, the activity sequences obtained for each resident were used to determine their actions. Regarding movement, activities that require moving through city streets were coded using the A-star algorithm (Tjiharjadi et al., 2022), which is a path finding algorithm in a graph that finds the costless route from a starting point to a destination. It was developed here, using an existing code (Suriabe, 2017), to formalise the population movements in the graph of the city's streets (i.e., without crossing walls and houses).

3 RESULTS

3.1 Cartography

A detail of the cartography that was obtained and which serves as support for the simulator is presented in Figure 1. The entirety of the domain (Figure 2) represents the urban core as a grid of 586 x 599 (351,014) cells arranged into 1,326 units or landplots (room, wall, door, shop, fence, yard, corridor, road, etc.). When food stocks have been identified during the census, they are reified and placed in the simulator as objects.

![Figure 1: Detail (31x24 cells, 0.21% of the domain) of downtown Dodel and the digitization of habitations (one cell = 1 m²).](image)
Nonetheless, it is mainly with the exhaustive census of human activities transcribed in the simulator that it was possible to obtain the digital twin that mimics diel activity in the city. Indeed, by combining detailed cartography with a faithful representation of human activity, we obtain the dynamic environment that organisms (ticks, mice, cats) must account for in their own activity.

Figure 2 for instance depicts a snapshot of the early morning activity recorded in the studied area. At that time, children leave for secular school in the south of the village (a), while others are already present at the Koranic school (b); the central market begins to bustle (c) with vendors already present in the shops surrounding the central market (d).

Using this representation, it is possible to simulate an environment that is close to reality, with periods of calm (meals, naps, rest) interspersed with areas that are occasionally active (school, mosque, central market, etc.). As animal agents (primarily mice) evolve in response to human activity, the patterns of activity resulting from their interaction (flee, hide)
with human presence appear realistic; for instance, there are significantly more mouse movements in the village during the middle of the night, even if movements of mice can still occur occasionally during the day, which is indeed observed. These movements are therefore not programmed here but generated by human activity.

### 3.3 Integrated Simulation

Field observations related to the presence of ticks, cats, house mice were analysed and reported on a static data portal (http://simmastbo.org/infos/052). Observed individuals were coded in the simulator. Within one standard scenario, the simulated system is composed of 489 humans, 135 cats, 175 mice, 68 ticks.

The simulator then makes it possible to specifically examine the various configurations obtained within the dwellings and the resulting interactions between agents within rooms. For instance (Figure 3), it is possible to study how inhabitants move from the courtyard to their room according to the time of day, how house mice hide when human activity falls into the categories "meal", "movement" and "awake" and how they are free to roam outside their nests when humans are distant or in the "rest" category.

![Figure 3: Detail (44x53 cells or m², 0.66% of the domain) of a dwelling including a courtyard and rooms where people live, as well as food-selling shops on the side of the road, and in this instance a pregnant female and an adult tick (data from observation and transcribed in the simulator).](image)

#### 3.4 Detailed Processes Simulation

Finally, the interactions between animals themselves, such as the reproductive behaviour of house mice (Figure 4) or the conditions under which ticks can infect or become infected from mice, can also be studied in detail (Figure 5).

![Figure 4: Study of behaviours and interactions between house mice within rooms (agents' label: agent id / target or desire in the absence of target): in this sequence, a male (4439) and a female (4436) mouse walk close to a room walls, the male then perceives the female, identifies it as its target and moves towards it. In the following sequence, reproduction can occur, and then the male and pregnant female resume their independent foraging activity.](image)

![Figure 5: Study of behaviours related to the interaction between vectors and reservoirs. In the example illustrated, the modelled room contains a mouse nest (704007) where ticks seek to feed on the mice present. Tick 704084 remains attached to mouse 13284 as it leaves the nest to feed, while tick 704082 remains in the nest.](image)
4 DISCUSSION

As part of the development of a multidisciplinary epidemic risk observatory, we have incrementally developed a digital twin of the Senegalese village of Dodel. In the context of successive stages, we conducted round trips between the field and digital coding to primarily (i) develop a high-resolution map of all homes and shared places in the city centre, (ii) introduce the census of all inhabitants, (iii) sample and formalise the presence of agents such as house mice, cats, ticks and bacteria and finally (iv) identify and code the details of the daily activity of all inhabitants. The simulator obtained is hence capable of representing, both globally at the level of the city centre and locally at the level of rooms, most of the interactions that can occur in this rural municipality.

The simulator was positioned at the core of the observatory’s operation. At the conclusion of the process, this approach proved fruitful by contributing three useful and complementary functionalities (Figure 6). Following the initial approach, the simulator constitutes a dynamic and integrated database of knowledge gathered by all involved disciplines. By providing agents with behaviours, not only can the entire system be represented in a unified manner, but it can also be brought into interaction. On the other hand, as the simulator is based on a bio-inspired approach, it is potentially robust to any improvement even not anticipated during the model’s initial design. For instance, in recent surveys, additional data were collected at the health post level on humans with fever, including information about the nature of the fever. Due to the bio-inspired approach used to develop the digital twin, such features can be added and incorporated into the simulator without calling the model into question.

1. Digital twins have been recognized as platforms for consensus building among stakeholders (Okita et al., 2019) which may be the case here for the simulator produced. Thanks to the high level of realism provided by the model following the digital twin approach, the dynamics generated are straightforward, easily comprehended, and communicable (movement of people, behaviour of mice in rooms and shops, etc.). In a logic known as companion modelling (Barreteau et al., 2003), the population and local (mayors, village chiefs) and regional (department) decision-makers can evaluate or criticize the results and help propose operational adjustments to the observatory in response.
2. As the simulator is based on a data-driven logic, it may exhibit issues during simulations that can be analysed to reveal areas or spaces where the field effort and the collected data are either insufficient or excessive. Within the framework of a constant back-and-forth between the field and the simulator's representation of the results, the tool constituted by the simulator is therefore a very useful instrument for guiding the field survey protocols.

4.1 Calibration Issues

The resulting model incorporates a large number of distinct actors, each with their own distinct deliberation and action behaviours. The urban environment presents additional challenges because agents must move in a variety of ways: along walls, in a network of streets, or linearly in rooms; avoiding numerous obstacles; failing to perceive nearby objects because walls obscure them, etc. On the other hand, field-based datasets are abundant, particularly those pertaining to human activity. This complexity presents challenges for assembly calibration, which can be linked to (i) data coding (e.g., a person sleeping in an awkward position in the middle of the street) or (ii) modelled behaviour (e.g., a hungry cat eats another cat). In this second category, the object-oriented and bio-inspired approaches used to develop the model allow for the resolution of the vast majority of inconsistencies, but leave room for a few incoherent behaviours. The identification of aberrant behaviours is more often than not based on the observation of simulations that are then corrected. This is a lengthy process, and the calibration of the model in this study is still ongoing but it does not constitute an insurmountable obstacle.

5 CONCLUSION

The process of creating a digital twin of a complex urban environment is lengthy and delicate. Nonetheless, once the computer system has stabilised, one can get a tool with decisive advantages. The first is to benefit from a dynamic restitution of knowledge that goes beyond what a sophisticated multidisciplinary DBMS could provide. The second is the digital twin's utility as a means of communication. As an anecdote, during the last restitution to the village authorities, the interest generated by the transmission of the results prompted the village chief to broadcast a message via the mosque asking for a good reception of scientists by the population as well as a request to the project team to make a public restitution of the work conducted to the population.

This work hence suggests that the digital twin paradigm can be adapted and applied to complex social issues such as the management of an observatory for the monitoring of epidemic risk involving multiple human-animal actors evolving in a variety of environments. This approach may therefore be proposed as an efficient mean to meet the methodological requirements raised by the EcoHealth approach.

ACKNOWLEDGEMENTS

The study was supported by the French National Research Institute for Sustainable Development (IRD), the ‘Centre de Biologie pour la Gestion des Populations’ (CBGP, UMR no.22 INRAe / IRD / Cirad / Supagro), the Cerise (grant IRD-FRB no.AAP-SCEN -20B III) and CEAMITIC (Univ. G.Berger, Saint-Louis Sénégal) projects as well as the ObsMice (West African Small Mammal Observatory Indicators of Environmental Change) network. We also wish to thank Y.Niang, M.Kane, N.Sarr, O.Sall, R.Dia, G.Diatta for their decisive field support and O.Sall for detailed mapping of the Dodel town. Finally, we would like to warmly thank the population and administration of Dodel for their hospitality and willingness to answer questions during our numerous surveys.

REFERENCES


