

Towards Virtual Epidemiology: An Agent-Based Approach to the Modeling of H5N1 Propagation and Persistence in North-Vietnam

Edouard Amouroux^{1,2}, Stéphanie Desvaux^{3,4}, and Alexis Drogoul^{1,2}

¹ IRD UR079 GEODES, 32 Avenue Henri Varagnat, Bondy - France

² Equipe MSI, IFI, 42 Ta Quang Buu, Ha Noi - Viet Nam

³ CIRAD, Campus International de Baillarguet F-34398 Montpellier - France

⁴ PRISE Consortium in Vietnam c/° NIVR, 86 Truong Chinh, Ha Noi - Viet Nam
{edouard.amouroux, alexis.drogoul}@ird.fr,
stephanie.desvaux@cirad.fr

Abstract. In this paper we claim that a combination of an agent-based model and a SIG-based environmental model can act as a “virtual laboratory” for epidemiology. Following the needs expressed by epidemiologists studying micro-scale dynamics of avian influenza in Vietnam, and after a review of the epidemiological models proposed so far, we present our model, built on top of the GAMA platform, and explain how it can be adapted to the epidemiologists’ requirements. One notable contribution of this work is to treat the environment, together with the social structure and the animals’ behaviors, as a first-class citizen in the model, allowing epidemiologists to consider heterogeneous micro and macro factors in their exploration of the causes of the epidemic.

Keywords: Multi-Agent Systems, Agent-Based Models, epidemiological models, environmental models, GAMA platform.

1 Introduction

Over the past few years, avian influenza spread from Asia to Europe and parts of Africa. Following this proliferation, a certain downturn has been observed thanks to concerted measures (improving hygienic practices, vaccination programs, etc). Yet, total eradication remains elusive. This disease remains a major threat to both the economy and public health. Within this context, the challenge for current epidemiology is to eradicate the virus, which requires understanding the various factors that may impact the proliferation of this infection. In particular, the local propagation of the virus and its re-emerging mechanisms are yet to be fully understood. Current hypotheses point an accusing finger towards: the presence of wild birds traditional farming practices and trading activities. None of these have been validated as of yet. Recent studies [1] suggest that human activities, particularly within the agricultural system, in dynamic interaction with the environment, play a key role in the spread of the disease at local levels (within a given district). To have a better understanding of these mechanisms, one of the possible methods is for epidemiologists to consider

simulated models of reality and to test hypotheses concerning the environment, the social structure, the behaviors of the birds, etc. Unfortunately, existing modeling techniques have not reached a level of maturity appropriate for such an exploratory use of simulation. The aim of this paper is to present an Agent-Based Modeling (ABM) approach, while detailing how it can be used to help epidemiologists answer their questions. After presenting the context and requirements of this work, we will review existing epidemiological models, explain why they do not fit the requirements and why we have chosen to develop our own tools on top of a multi-agent platform.

2 The Context of Avian Influenza in Vietnam

Avian influenza (HPAI) epidemics occurred recurrently in Vietnam since mid-2003. While the outbreaks in both North and South Vietnam are similar the underlying variables of their respective poultry production industries are different: climactic and environmental conditions, variety of circulating virus strain [2] and poultry production organization [3] are distinctive in each region.

2.1 Two Epidemiological Questions to Address

As of today, many basic epidemiologic questions (about avian influenza) still need to be answered. At a macro level, propagation is not precisely understood but the general tendencies are. At the micro scale (e.g. village or commune level), the explanations of propagation mechanisms are hazy at best. Understanding what is happening at these micro-levels is probably the key to controlling the epidemic. Hence many investigations are focusing on them, particularly on these key questions:

- (1) **Propagation mechanisms:** how does the virus propagate locally (from birds-to-birds, humans, environments, farms, markets, etc.)?
- (2) **Persistence mechanisms:** how can the virus re-emerge in a previously uninfected area, sometimes months after the end of the previous wave of the epidemic?

2.1.1 Local Propagation Mechanisms

Both mechanisms can be explained by the presence of wild birds [4] and by trading activities. Other recent hypotheses concern the role of the agro-system. Recent policies focus on small farmers for their lack of biosecurity. Studies are also concentrating on semi-industrialized farm—as they have a much wider span of influence—yet biosecurity is low. As fully industrialized farms have adopted modern, state-of-the-art biosecurity systems, they present a lesser concern for the experts. Whatever the considered production sector is, the entire agro-system (the farm, suppliers and customers) is being monitored extensively by epidemiologists.

2.1.2 Re-emergence, or Persistence Mechanisms

Re-emergence is another concern. The prevailing hypothesis is that wild birds act as reservoirs but validation is lacking at present [4]. Another hypothesis concerns the environment itself as a reservoir. In-vitro experiments are currently being carried out on the persistence of H5N1 within the environment and especially in water [5]. It has been demonstrated that avian influenza viruses are likely to remain infective several months while in places like ponds [5].

2.2 Addressing Epidemiological Questions through Simulated Experiments

2.2.1 Towards Simulated Experiments

Local-scale propagation, persistence and re-emergence processes are a complex matter where many actors (birds, domestic or not, environment, human activity, etc...) intervene. Unfortunately, the data that has been produced to date is insufficient in term of completeness and reliability. This situation will probably remain unchanged in the near future. To have more reliable data, in-vitro studies are being conducted, but the conditions are too far from reality to be easily transposed to field reality. Since neither field nor in-vitro studies can be completely satisfactory, experts turn towards models and simulation experiments to validate their hypotheses, and explore new ones.

2.2.2 Requirements for "Virtual Epidemiology" Experiments

As we will see in the next section, existing modeling techniques are not really suited to such "in silico" experiments. Indeed, the ideal tool would need to respond to several requirements expressed by epidemiologists. The most important are listed below:

- (1) The *environment* should be represented *extensively* together with its *own dynamics*. This representation should also enable the use of various and heterogeneous descriptions (geographical, social, ecological, etc).
- (2) Epidemiologists need the ability to work both at the *population level* and directly on models of *individuals*, depending on the hypotheses requiring validation.
- (3) *Interactions at the population or individual levels*, but also *with the environment*, should be modeled explicitly, since a change in these interactions may have an impact on the dynamics of the epidemic.
- (4) These data are usually expressed using *heterogeneous formalisms* that the expert needs *to be able to reuse* in the modeling of the system.
- (5) The simulation platform should be able to recreate an environment in which "*virtual experiments*" could be run (exploration of parameters, etc.)

3 Epidemiological Models: A Brief Review

In this section, we review the existing epidemiologic models in the light of these requirements. Although the literature on the subject is quite important, we will see that most of the offers do not fulfill them completely yet.

Table 1. Summary of the relationships between questions and models

How to ...	Model	Referece
Represent and study epidemics at a global level?	SIR Bayesian Networks HMC	Yorke 79 Abbas 04 Durand 99
Account for the heterogeneity of the population?	Micro-simulation	Artzrouni 01
Represent contact patterns?	Cellular automata Social networks	Turner 01 Ghani 07
Represent different individual behaviors?	Individual based	Wilensky 98
Study the role of the environment?	Agent based Agent based + GIS	Muller 04 Badariotti 05
Enhance the representation of the environment?	Agent based + GIS	Badariotti 05

Although many types of model have been proposed and used for epidemiological ends, every one of them still lacks the ability to address some of the requirements expressed by epidemiologists in this type of research (see 2.2.2).

It is plain that global models such as SIR [6], bayesian networks [7] or hidden Markov Chain [8] will not work when considering small-scale situations. Micro-simulation, [9] & [10] will work, but lacks the ability to take situatedness into account. This can be addressed by cellular automata [11] or social networks [12], depending on the level of complexity to be modeled. These models cannot represent fully heterogeneous agents (internal state + behaviors)—while ABM [13] does bring this capacity to the table. Finally, this kind of model generally lacks a detailed expression of the environment—among our top concerns in this application. To address this issue, computer scientists, [14] & [15], have proposed the joint use of ABM and GIS but they have not yet fully implemented this idea.

Thus, the natural choice for us, was to follow an agent-based approach, coupled with a detailed and flexible representation of the environment based on a combination of grids and GIS. Additionally, we decided to make the environment (and its components) a “first-class citizen” of the model, provided with its own attributes and behaviors, rather than just considering it as a topological surface. This extension to “classical ABMs” will be described in the next section through the presentation of the HPAI in North Vietnam model built in cooperation with epidemiologists.

4 Conceptual Model of Avian Influenza Propagation

In this section, we introduce the context of this research, followed by the representation of the environment and actors. All the choices made in accordance with epidemiologists and field specialists will be explained.

4.1 Frame of the Epidemiologic Study

The study takes place in North Vietnam, where epidemiologists focus on “local” mechanisms, i.e. mechanisms that occur at a scale comprised between the village and the district levels (around 50 km²). The model is then geographically limited by the bounds of a province (a few hundreds km²). The main geographical entity is the village, which is considered by epidemiologists as a “coherent epidemiologic unit”. Their hypothesis, present in the model, is that communes and districts are not really relevant to consider when it comes to studying the local causes of propagation. This assumption may however be easily revised in future occurrences of the model. As the environment may be a reservoir for the virus, we will consider that every place may allow virus survival. In addition all entities in the model can be also infective.

The “village” environmental unit

A “traditional” Vietnamese village consists of an inner-village space with a main street, a few dozen to more than a hundred households with some poultry (traditional farms) and pets. This inner space is surrounded by rice fields, watered lands (which enter the village), other cultures and is protected from the flooding by a dike. The organization of the village is described in [16] while statistic will be provided by the Vietnamese census and longitudinal surveys currently conducted.

The environment

The surroundings of the village consist of agricultural and “natural” lands (river, forest, etc.). As we do not focus on long distance propagation, we consider this distant space as homogeneous. Conversely, the inner-village and agricultural lands are represented in detail, especially the possibility that the latter may act as a reservoir for the virus. According to epidemiologists, viral dynamics within the environment are susceptible to: *altitude, Ph, temperature, solar exposure, level of organic matter* present in water, *watered or not*. These parameters define the ecological dynamics of the system. These parameters vary greatly in natural or cultured land we can consider them to be static in the inner village and roads.

4.2 Relevant Actors

Actors of the pathogen system act at different scales and are organized in 2 main structures: the village and the poultry production chain. The other levels of organization (like administrative levels) are neglected in the model, as they do not impact local scale dynamics.

Actors of the Traditional Village

In the village important actors are organized around the farm. The type and dynamics of both are defined according to the production type and the production sector (see below).

Farm and markets

We can classify farms by their size and production techniques in four sectors according to [17]:

- *Traditional farming*: mixed poultry, local scale interactions, no bio-security at all.
- *Semi-industrial & industrial*: targeted production, district to province size scale interactions, low/medium bio-security level (medium/high for industrial farms).
- *Fully integrated*: targeted production, province size scale interactions, sophisticated bio-security systems (thus, not considered in the model).

The production type and sector determine most of the farm’s characteristics in terms of *herd (size, lots organization, vaccination coverage, species and breeds)*, the type of premises (caged, fully confined, pond), which impacts the *bio-security level*, etc. They also determine the dynamics of the farm and its acquaintances (processing chain).

Markets are also a key location within the village. It acts both as a reservoir for the virus and an exchange place, as there is extensive contact between processed and living poultry. The latter can be brought back home newly infected if not sold.

Poultry

According to experts, all the poultry within a flock are very similar in terms of behaviors (gregarious animals), and characteristics (homogeneous lots). So we can aggregate these individuals and represent only the poultry flock. This, plus homogeneous mixing occurring within the flock allow us to represent the virus transmission within the flock with a SIR model. A traditional farm’s flock will be addressed by developing an adapted SIR model as they are not homogeneous lots. The flock behavior depends on the farm type and sector, the poultry can be confined, access an adjoining or distant secured zone (a pond, channel, etc), or be free ranged.

Humans

Humans may play the role of mechanical vector in local propagations. However, their impact is considered as very low when compared to poultry flocks. Consequently, we neglect them in this first model. (Para-) veterinarians are a different case, as they travel among numerous farms and are not always fully trained they may impact the disease dynamic, thus they will be represented explicitly, going farms to farms.

Animals

Wild birds, which are considered by most epidemiologic surveys as not playing a noteworthy role in local propagation, are not to be considered. Peridomestic birds, farm animals, pets and fighting cocks are also removed from this model, as the data available on them regarding their role in avian influenza is not clear though surveys are being conducted and their integration is planned.

4.2.1 Outside the Traditional Village

Villages are also linked to poultry production networks, which have a much wider scale dynamic and a possible strong impact because they connect all the farms and may constitute a good propagation system.

Mainly, we can consider ([3] and expert knowledge) four actors of this production chain: traders (highly variable in term of size), transporters, slaughterhouses and selling points. Traders can stock, trade among themselves and manage their transporters, who carry living or slaughtered poultry from one chain production node to another.

Slaughterhouses “just” process the poultry and thus have no relevant dynamic. The only point of sale of processed poultry (i.e. supermarket) can be neglected as it is considered as a disease end point (no infection possible). On the contrary, local, district or provincial markets need to be represented in the same way as the commune.

We have a fairly complex system with both a detailed environment and numerous actors to represent—but, at this point, the general structure has been defined. The environment is split into three main categories: a static inner village, the agricultural surroundings (which is dynamic) and the far surrounding areas (considered static and homogeneous). Actors are village-related and connecting villages are related.

As all needed data are not yet available, the creation process of the model will be incremental. We are conscious that we must remain prepared to integrate any entities that may be declared relevant from surveys being conducted presently.

4.3 Implementation

To implement this model we use a versatile ABM simulation platform, GAMA [18]. The specific feature of this platform is that it allows modelers to work with an artificial data environment (a grid, for instance) in the same way as field data environments (i.e GIS) and provide for the seamless integration of both if needed.

4.3.1 The Environment

Similarly to [15] we use a mix of grids, unavailable data and within-simulation generated data, and GIS objects representing field data information. Both contain the environmental data as described in the previous section. These parameters are used to define the endogenous dynamic of the environment, and the virus’ evolution in

particular. This allows epidemiologists to work on real data (GIS) and incomplete data but also to infer unavailable data and test the plausibility of such data.

4.3.2 Agents

Classically, we use reactive agents but they are situated in a complex environment (GIS & grids) and they have also persistent action (for example the daily behavior of a poultry flock). The modeling of poultry is of particular interest. A flock is homogeneous in term of structure and behaviors thus it can be represented with one agent representing the group. Although, they differ in regards to virologic history, as they are not infected all at the same time and respond differently to the disease. Consequently, we added a matrix of individuals in each flock as an agent, in the manner of micro simulation.

5 Conclusion

In this paper we have presented a brief review of epidemiologic questions and the models adept at responding to those questions. We moved away from the field of mathematics, while addressing global-scale questions, using computer-based models and especially IBM which are better suited for a small-scale problematic. Afterwards, we presented a model to study small-scale epidemiological phenomenon in the context of avian influenza in North Vietnam: local propagation and re-emergence mechanisms. This model differs from those previously proposed, as it treats the environment as a first-class citizen of the system. Here the environment is not only a topological surface, it can contain heterogeneous data, have its own dynamic and can be multiple (GIS + grid in our application). In future works, validation is our most important concern. Although this model is explanatory and exploratory rather than predictive in nature, we consider validating the predictive capabilities of the model paramount to our future work.

References

1. Gilbert, M., Xiao, X., Pfeiffer, D.U., Epprecht, M., Boles, et al.: Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proc. Natl. Acad. Sci. USA.* 105, 4769–4774
2. Dung Nguyen, T., Vinh Nguyen, T., Vijaykrishna, D., Webster, R.G., Guan, Y., Peiris, J.S.M., Smith, G.J.D.: Multiple sublineages of influenza A virus (H5N1), Vietnam, 2005–2007. *Emerging Infectious Diseases*, 14, 4 (2008)
3. Agrifood consulting international: *The Impact of Avian Influenza on Poultry Sector Restructuring and its Socio-economic Effects* (2006)
4. Chen, H., Li, Y., Li, Z., Shi, J., Shinya, K., Deng, G., et al.: Properties and dissemination of H5N1 Viruses isolated during an influenza outbreak in migratory waterfowl in Western China. *J. of Viro.* 80, 5976–5983 (2006)
5. Ito, T., Okazaki, K., Kawaoka, Y., Takada, A., Webster, R.G., Kida, H.: Perpetuation of influenza A viruses in Alaskan waterfowl reservoirs. *Arch. of Viro.* 140(7), 1163–1172 (1995)

6. McKendrick, A.G.: Applications of mathematics to medical problems. Proc. of the Edinburgh Mathematical Society 44, 1–34 (1925)
7. Abbas, K., Mikler, A., Ramezani, A., Meneze, S.: Computational epidemiology: Bayesian disease surveillance. In: Proc. of ICBA 2004 proceedings (2004)
8. Durand, B., Mahul, O.: An extended state-transition model for foot-and-mouth disease epidemics in France. Preventive Veterinary Medicine 47, 1–2 (2000)
9. Artzrouni, M., Gouteux, J.P., et al.: Population dynamics of sleeping sickness: A micro simulation. Simulation & Gaming 32(2), 215–227 (2001)
10. Brouwers, L.: MicroPox: a Large-scale and Spatially Explicit Microsimulation Model for Smallpox Transmission. In: Proc. of the Intl. Conf. of Health Sciences Simulation (2005)
11. Turner, J., Begon, M., Bowers, R.: Modelling pathogen transmission: the interrelationship between local and global approaches. Proc. Biol. Sci. 270(1510), 105–112 (2003)
12. Ghani, A.: Models for an Outbreak of Avian Influenza in the UK Poultry Flock, Journées de modélisation en épidémiologie animale du CIRAD (2007)
13. Wilensky, U.: NetLogo Virus model (1998), <http://ccl.northwestern.edu/netlogo/models/Virus>
14. Muller, G., Grébaud, P., Gouteux, J.P.: An agent-based model of sleeping sickness: simulation trials of a forest focus in southern Cameroon, Éd. Sc. et Méd. Elsevier, Amsterdam (2004)
15. Badariotti, D., Banos, A., Laperrière, V.: Vers une approche individu-centrée pour modéliser et simuler l'expression spatiale d'une maladie transmissible: la peste à Madagascar, cyberge (2007), <http://www.cyberge.eu/index9052.html>
16. Bénédicte, S.J.: Atlas Bac Hung Hai, Gret (personnal communication) (1999)
17. Desvaux, S., Vu, D.-T.: A general review and description of the poultry production in Vietnam. Agricultural publishing house (2008)
18. Amouroux, E., Chu, T.Q., Boucher, A., Drogoul, A.: GAMA: an environment for implementing and running spatially explicit multi-agent simulations. LNCS (LNAI). Springer, Heidelberg (2007)