A DBN model to study the influence of epidemio surveillance networks on phytosanitary treatments

Jean-Noel Aubertot, Marie-Josée Cros, Sabrina Gaba, Nathalie Peyrard, Xavier Reboud, Régis Sabbadin

A DBN model to study the influence of epidemio surveillance networks on phytosanitary treatments intensity

Jean-Noël Aubertot [1], **Marie-Josée Cros** [2], **Sabrina Gaba** [3], **Nathalie Peyrard** [2], **Régis Sabbadin** [2], **Xavier Reboud** [3]

1. *AGIR - UMR 1248, INRA/INPT, Toulouse*
jean-noel.aubertot@inra.fr
2. *MIAT - UR 875, INRA, Toulouse*
marie-josee.cros/nathalie.peyrard/regis.sabbadin@inra.fr
3. *Agroécologie - UMR 1347, INRA/AgroSup/Université de Bourgogne, Dijon*
sabrina.gaba/xavier.reboud@inra.fr

RÉSUMÉ. *Le rôle des Réseaux d'Epidémio Surveillance (RES) pour le contrôle des épidémies dans les cultures est de plus en plus important. Cependant, l'influence de l'information fournie par le RES sur le nombre de traitements phytosanitaires appliqués et la taille de l'épidémie est mal connu. Nous proposons un modèle de DBN pour représenter la dynamique d'un bioagresseur dans un parcellaire, afin d'explorer l'influence de la taille spatiale et temporelle du RES sur ces critères. Une analyse par simulation montre que pour les adventices et les pathogènes du sol, il est possible de réduire le nombre de traitements significativement, pour une augmentation modérée de la taille de l'épidémie, en augmentant la taille du RES. Cela illustre comment un DBN peut être utilisé pour identifier des leviers pour la réduction de l'usage des pesticides.*

ABSTRACT. *The role of Epidemiological Surveillance Networks (ESN) is becoming more and more important to control epidemics in agricultural areas. The influence of the ESN outputs on the number of phytosanitary treatments applied, and on the epidemic size is not well understood. We used a dynamic Bayesian network model of a pest dynamics to explore the influence of spatial and temporal sizes of the network on these criteria. A simulation analysis showed that for weeds and soil-borne diseases, by increasing the ESN size, the number of treatments can be significatively reduced with a moderate increase of the epidemics size. This illustrates how DBN models may help in identifying ways to reduce pesticide use.*

MOTS-CLÉS : *Suivi de bioagresseur, règle de décision, simulation, analyse de sensibilité*

KEYWORDS: *Pest monitoring, decision rule, simulation, sensitivity analysis*

DOI:10.3166/RIA.?.1-8

1. Introduction

Globalisation, environmental and climate changes multiply the risks of emergence and re-emergence of diseases or animal pests on crops. Pest control is difficult because decision choices must be made early to ensure their efficiency under uncertain results of actions and uncertain system dynamics. To control and prevent epidemics, the role of epidemiological surveillance systems (Reboud *et al.*, 2017) is becoming more and more important. In France, the Ecophyto (EcoPhyto, 2015) national action plan to reduce pesticide use strengthened the role of the national Epidemio Surveillance Network (ESN).

An ESN is a set of fields in an agricultural area that are observed by experts able to detect an infection. This information is then provided to the farmers who can use it to decide whether they apply or not a phytosanitary treatment. So the ESN information could be used to limit unnecessary treatment actions. However, the influence of the spatial size (how many surveyed fields) and temporal size (the length of the history of observations) of the network on the number of treatments, the epidemics size and therefore the net margin is not well understood.

In this study we compare different ESN using simulations of a Dynamic Bayesian Network model (DBN, (Jensen, 2001)) of the spatio-temporal propagation of the pest throught a set of crop fields. We combine the model with a decision rule at the field level that combines the ESN information (shared by all farmers) and the local knowledge of the farmer on the field sanitary status history (private information) to decide whether to apply a treatment or not. Considered ESNs have different spatial and temporal sizes. Results are obtained for three parameterizations of the DBN model corresponding to three typical pest types: soil-borne pathogens, weeds and pest insects. We also present a sensitivity analysis on the model parameters.

2. DBN model of pest dynamics in crop fields

2.1. Model when treatment actions are known

Let us consider a set of n crop fields in a landscape, that can be affected by a pest. We denote by X_i^t the state of field i at time t: 0 for non infected and 1 for infected. The pest can spread from field i to a neighbour field j, and N_i is the set of indices of the neighbour fields of i. At each time step, a decision to apply a phytosanitary treatment can be made. This is represented by the binary variable A_i^t, equal to 1 if a treatment is applied and zero otherwise. If $X^t = (X_1^t, \ldots, X_n^t)$ is the sanitary status of all fields at time t and $A^t = (A_1^t, \ldots, A_n^t)$ are the actions chosen at every field at time t, we model the temporal distribution of X^t given A^t as a Dynamic Bayesian Network: at the landscape level the transition probability is

$$P(X^t|X^{t-1}, A^t) = \prod_{i=1}^{n} P(X_i^t|X_i^{t-1}, X_{N_i}^{t-1}, A_i^t).$$

Then the individual transition probabilities are modeled as in a per contact propagation process where a field can become infected either by colonization from neighboring fields or by long distance propagation. If $I_i^t = \sum_{j \in N_i} X_j^t$ is the number of infected fields in the neighborhood of field i the transitions are:

$$P(X_i^t = 1 \mid X_i^{t-1} = 0, X_{N_i}^{t-1}, A_i^t = 0) = \epsilon + (1-\epsilon)(1-(1-\rho)^{I_i^{t-1}}) = P_{01}(I_i^{t-1})$$
$$P(X_i^t = 1 \mid X_i^{t-1} = 0, X_{N_i}^{t-1}, A_i^t = 1) = (1-\gamma)P_{01}(I_i^{t-1})$$
$$P(X_i^t = 1 \mid X_i^{t-1} = 1, X_{N_i}^{t-1}, A_i^t = 0) = \nu + (1-\nu)P_{01}(I_i^{t-1}) = P_{11}(I_i^{t-1})$$
$$P(X_i^t = 1 \mid X_i^{t-1} = 1, X_{N_i}^{t-1}, A_i^t = 1) = (1-\gamma)P_{11}(I_i^{t-1})$$

The model parameters are ϵ, the probability of success of long distance propagation of the pest, ρ, the probability of success of colonisation from a neighbor field, ν, the probability of pest survival if the field is not treated, and γ the probability of success of the phytosanitary treatment.

2.2. Decision rule for treatment

In the model, the states of some fields are observed each year, to inform the farmer who must decide if he applies a phythosanitary treatment or not. These fields are the same every year and form the Epidemio Surveillance Network (ESN). The temporal horizon of the ESN is denoted h_{ESN}, it means that at year t the information is available back to year $max(1, t - h_{ESN})$. We denote by O the set of indices of the fields in the ESN and we define $p_i^t(ESN)$, the probability that field i is infected at time t given the knowledge of the states of the fields in the ESN:

$$p_i^t(ESN) = P(X_i^t = 1 \mid X_O^t = x_O^t, \ldots, X_O^{max(1,t-h_{ESN})} = x_O^{max(1,t-h_{ESN})})$$

We model the pest management strategy at time t in field i as a decision rule at the field level that combines two pieces of information: (i) the infection probability $p_i^t(ESN)$ of the field i estimated from the ESN observations and the DBN model (shared information), and (ii) the sanitary state $X_i^{max(t-h_{loc},1)}, \ldots, X_i^{t-1}$ of field i during the h_{loc} preceeding years (private information), where h_{loc} is the temporal horizon of the local information.

Let us first assume that the only information available for deciding is $p_i^t(ESN)$. We can use this information to derive a decision rule for treatment which aims at maximizing the expected net margin of the field at time t. Treatment is chosen as soon as the expected net margin with treatment is larger than the expected net margin without treatment. The expected net margin without treatment, m, depends on the yield ($yield$), the price of crop ($price$), the total cost of production (c), the cost of phytosanitary products (c_{phyto}) and q the proportion of yield preserved in presence of the pest. It is equal to $(1 - p_i^t(ESN))W + p_i^t(ESN)w$ with $W = price.yield - (c - c_{phyto})$ and $w = q.W$. The expected net margin with treatment, m_{phyto} is equal to

$(1 - p_i^t(ESN))(W - c_{phyto}) + p_i^t(ESN)(\gamma W + (1-\gamma)w - c_{phyto})$. Then above a threshold value of $p_i^t(ESN)$, we have $m < m_{phyto}$ and treatment is applied: $A_i^t = 1$.

To take into account the private information available to the farmer in field i, we assume that the farmer updates this threshold with the knowledge he has of the history of his field $(X_i^{t-1}, X_i^{t-2}, \ldots)$, leading to a lower threshold (meaning more treatment decisions) if the field has often been infected, and a larger threshold if the field has never been infected (details of this update are omitted for sake of brievity). It is not necessary to include $A_i^{t'-1}$ for $t' < t$ in this private information since X_i^{t-1} is known. In addition, we assume each field belongs to a different farmer so actions in others fields that i are not known by farmer i.

3. Design of the simulation study

Criteria for comparing ESN. The objective of an ESN is to provide information to the farmer. A farmer who would have decided to preventively treat the field against a potential pest invasion, can decide not to treat if the ESN information indicates a low incidence of the pest in the area. The opposite can also happen. So, different ESN can be compared in their capacity to control the pest invasion and limit the treatments. So the two first criteria are the number of infected fields and the number of treatments applied. They both contribute to the third criterion that we will consider: the farmer net margin.

ESN characteristics. We compared eight ESNs on a grid of 144 (12 by 12) crop fields: four spatial sizes (1%, 10%, 25%, 50% of all fields, uniformly spatially distributed, stable in time, see Figure 1) and two temporal sizes (preceding year $h_{ESN} = 1$ or all available preceding years ESN observations $h_{ESN} = H$), and $h_{loc} = 1$.

Pest and agronomic parameters. Three types of crop pests were considered: soil-borne pathogens, weeds, pest insects. If a pest is in a field, it can colonize the 4 closest fields. These types have contrasted spatial dispersion and temporal persistence parameters (see Table 1, established from expert knowledge).

Tableau 1. Values of DBN parameters for the three pest types

pest type	ϵ	ρ	ν
soil-borne pathogens	0.05	0.10	0.50
weeds	0.15	0.20	0.50
pest insects	0.30	0.40	0.25

We considered oilseed rape fields. Parameters associated to crops are yields for healthy or injured crops, oilseed rape price, production and treatment costs, treatments efficacy. Due to the difficulty to collect data at the spatial and temporal scale considered, we did not estimate these parameters, they were set by expertise and studies in France (FOP, 2015 ; INOSYS, 2014 ; Cros *et al.*, 2017).

Inference of $p_i^t(ESN)$. When simulating a trajectory of the DBN model, at each time step it is necessary to compute $p_i^t(ESN)$ for each field i in order to determine the action prescribed by the decision rule. Simulation of trajectories is time consuming without a fast approximate inference method to compute the $p_i^t(ESN)$ at time t for all fields. Since the observations are the same for each field (the one provided by the ESN) it is possible to approximately compute them simultaneouslty for all field using Loopy Belief Propagation (LBP) or Gibbs Sampling (GS). We compare both methods using the BN toolbox (Murphy, 2001). Since we observed a bug after a few iterations when running LBP (both probabilities of pest presence and absence estimated to zero) we eventually used GS. Note that we are currently developping the Matlab GMtoolbox (http://www.inra.fr/mia/T/GMtoolbox) for inference in graphical models using generalized belief propagations algorithms. This will enable faster inference.

Simulation protocol. For each ESN, we considered three initial states for the infection (4 fields, either at the corner, in the middle of a border, or close to the center of the grid). For each initial state we ran 60 simulations of $H = 8$ years. Mean values of the 3 criteria were computed over years 5 to 8. For comparison purpose, in addition to the decision rule derived from the ESN, we also considered the two extreme strategies consisting in never treating and always treating. The implementation was made in Matlab. Code is available on Figshare (https://doi.org/10.6084/m9.figshare.4675759.v3).

Sensitivity analysis. Using sensitivity analysis, we studied the influence of the 6 parameters related to the type of pest: (ϵ, ρ, ν) for the pest dynamics and (c_{phyto}, γ, q) for the associated net margin. Since simulating the DBN model requires a high computer execution time, we used a kriging metamodel (function km of R package DiceKriging) to approximate the relationship between the ESN and the model outputs. We performed one sensitivity analysis per pest type, by zooming in the metamodel on an hypercube around the values of all pest parameters.

4. Results

4.1. Simulation analysis

Simulations for soil-borne diseases (see Figure 1) and weeds give similar results. The mean number of treatment applications decreases when spatial network size increases (for h=1, from -11 % to -48 % ; for h=8, from -12 % to -46%) and also when the historic length decreases (for RES1, equal ; for RES2, -1 % ; for RES3, -4 % ; for RES4, -2%), but with a stronger effect of the network spatial size.

Mean economic net margin increases when spatial and temporal network sizes increase, in particular for ESNs with larger spatial size and for weeds (because their annual treatment cost is much higher). Although infection intensity slightly increases with larger ESN (for h=1, from 18 % to 38 %; for h=8, from 16 % to 31 %), due to a decrease in pesticide use, soil-borne diseases and weeds remain almost at the same level than when a treatment is applied systematically.

The results for pest insects were quite different. A treatment decision was always chosen whatever the considered ESN due to the high dispersal ability of pest insects, implying a high risk of injuries, together with a low treatment cost.

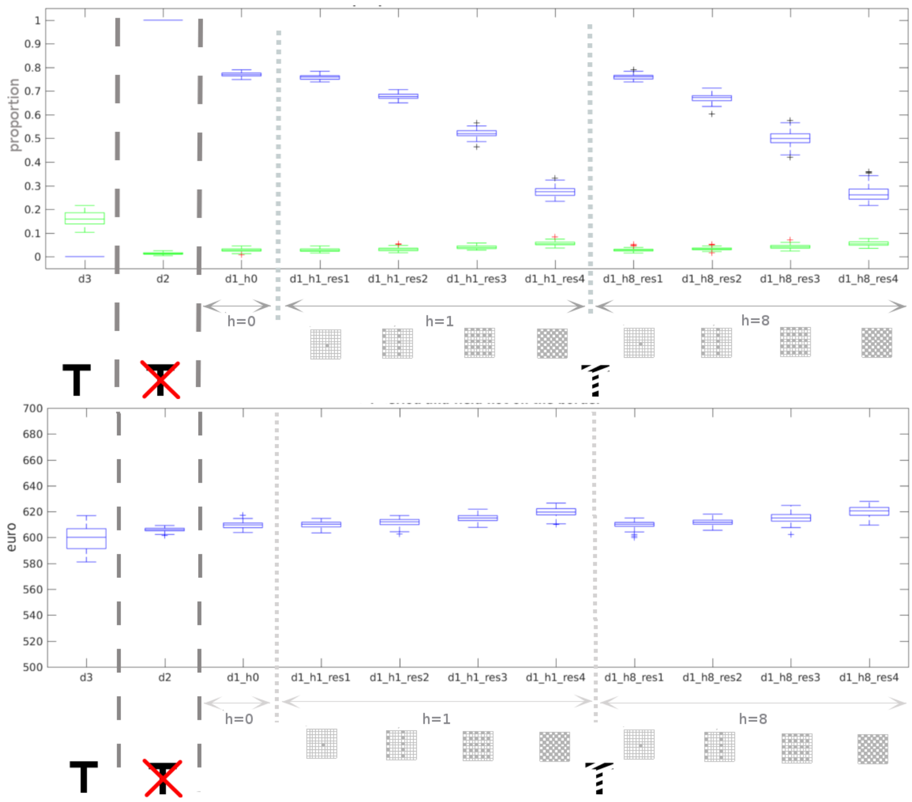

Figure 1. Influence of ESN temporal and spatial sizes. Top graph: mean percentage of treatment decision (blue) and mean percentage of infected fields (green). Bottom graph: mean economic net margin. On the horizontal axis, the considered strategies are: never treat (d3), always treat (d2), strategy only based on the private information (d1-h0), strategies combining ESN information and private information with $h = 1$ or $h = 8$ and ESN1 to ESN4 (d1-h1-res1 to d1-h8-res4).

4.2. Sensitivity analysis

For soil-borne pathogens and weeds, the more sensitive factors are annual cost of treatment (ct), percentage of yield preserved when infected (q) and probability of colonisation from a neighbouring field (ρ) (see Table 2). The first one is a socio-economic factor while the two other factors are related to natural caracteristics and agricultural practices.

For insect pest, the main sensitive factors are probability from neighbouring fields (ρ), yield when infection (q) and probability of treatment efficacy (γ) (see Table 2). In this case, factors related to natural characteristics and agricultural practices have more influence on the output variables than the socio-economic factor annual cost of treatment (ct).

Tableau 2. Sobol indices. Top: mean percentage of treatment, Middle: mean percentage of injured fields, Bottom: mean net margin. Darker cells indicate the larger Sobol indices.

Nb. treatments	ϵ	ρ	ν	ct	γ	q
Soil-borne pathogens	0.06	0.27	0.03	0.54	0.02	0.32
Weeds	0.05	0.28	0.03	0.53	0.02	0.31
Pest insects	0.10	0.21	0.02	0.16	0.48	0.69
Nb. infected fields	ϵ	ρ	ν	ct	γ	q
Soil-borne pathogens	0.04	0.03	0.01	0.41	0.15	0.41
Weeds	0.03	0.03	0.01	0.40	0.12	0.37
Pest insects	0.03	0.02	0.00	0.30	0.13	0.48
Net margin	ϵ	ρ	ν	ct	γ	q
Soil-borne pathogens	0.14	0.45	0.01	0.23	0.29	0.60
Weeds	0.14	0.45	0.01	0.23	0.30	0.57
Pest insects	0.07	0.86	0.08	0.00	0.00	0.00

5. Discussion

The main conclusions of this study are that the same ESN cannot be efficient for all types of pests. The second is that, for weeds and soil-borne diseases, by increasing only the ESN size, the number of phytosanitary treatments can be significatively reduced with a moderate increase of the epidemic size. For insects that can spread very fast, whatever the ESN size the decision is always to treat. This study illustrates how DBN models may help in identifying ways to reduce pesticide use.

The results presented are dependent of the way we have modeled the integration of the ESN and the private information in the decision rule. We have modeled it in two steps, first a threshold on $p_i^t(ESN)$ and then an update of the threshold using the private information in the field. Another choice would be to compute the probability of infection of field i given both information (from ESN and private) and to directly compute the threshold on this probability. However this would be computationaly more demanding since the evidence received at field i would be different from the evidence received at another field. Simultaneous computation of the conditional probabilities of infection using Gibbs Sampling would not be possible anymore.

Furthermore, the model could be used with different decision rules or extended to take into account ESN cost, agricultural practices that reduce biotic risks or increase treatment efficacy, landscape heterogeneity. It could also be used with different choices of neighbourhood to represent specific propagation situations, like wind orientation or agricultural machines trajectories (with the limit that increasing the neighbourhood size will increase the computational time). The model may then serve as an interactive tool for discussion between farmers and advisers to better understand the spatial and temporal connections underlying pest dynamics and help designing management strategies at the landscape level (Debaeke *et al.*, 2018).

Finally, in this study we chose to evaluate ESNs. Of course it would be interesting to go one step further, and to design an optimal ESN (in terms of infection size, number of treatments, ...). However this would require the optimization of a value function which is already hard to compute. This is an open question which is left for further research.

Acknowledgements. We thank Victor Picheny and Robert Faivre for their valuable support on the sensitivity analysis procedure. This work has been partially supported by the French Ecophyto project VESPA.

Bibliographie

Cros M.-J., Aubertot J.-N., Peyrard N., Sabbadin R. (2017). GMDPtoolbox: a Matlab library for designing spatial management policies. application to the long-term collective management of an airborne disease. *Plos ONE*, vol. 12, n° 10.

Debaeke P., Aubertot J.-N., Bardy M., Bertuzzi P., Cellier P., Constantin J. *et al.* (2018). Data collection and integration for landscape agronomy. *Landscape Agronomy, en révision*.

EcoPhyto. (2015). *Le gouvernement présente la nouvelle version du plan ecophyto.* Press Communication. Consulté sur http://agriculture.gouv.fr/sites/minagri/files/151026_cp_ecophyto.pdf

FOP. (2015). *Coûts de production, evolution des recettes en oléagineux, protéagineux et céréales.* Fédération française des producteurs d'oléagineux et de protéagineux, FOP. Consulté sur http://www.fopoleopro.com/etude-fop-couts-de-production-evolution-des-recettes-en-oleagineux-proteagineux-et-cereales-observatoire-couts-de-production-fop-cer/

INOSYS. (2014). *Systèmes grandes cultures en poitou-charentes, edition 2015 : campagne 2013 et conjoncture 2014.* Consulté sur http://www.deux-sevres.chambagri.fr/fileadmin/publication/CA79/17_Grandes_Cultures/Documents/INOSYS_CAS_TYPES_20132014GRCULT.pdf

Jensen F. V. (2001). *Bayesian networks and decision graphs.* Secaucus, NJ, USA, Springer-Verlag New York, Inc.

Murphy K. (2001). The Bayes Net toolbox for Matlab. *Computing science and statistics*, vol. 33, n° 2, p. 1024-1034. Consulté sur https://www.cs.utah.edu/~tch/notes/matlab/bnt/docs/bnt_pre_sf.html

Reboud X., Carpentier A., Aubertot J.-N., Lemarié S., Peyrard N., Turenne N. *et al.* (2017). Ce que l'épidémiosurveillance apporte ou n'apporte pas encore à la moindre dépendance aux produits phytosanitaires. *Innovations Agronomiques*, n° 59, p. 81–91.