



HAL
open science

Un modèle spatiotemporel pour analyser le commencement d'une épidémie, application aux rouilles du blé

Samuel S. Soubeyrand, Ivan I. Satche, Michael Hohle, Leonard Held

► **To cite this version:**

Samuel S. Soubeyrand, Ivan I. Satche, Michael Hohle, Leonard Held. Un modèle spatiotemporel pour analyser le commencement d'une épidémie, application aux rouilles du blé. Colloque Emergences, Oct 2006, Paris, France. hal-02751073

HAL Id: hal-02751073

<https://hal.inrae.fr/hal-02751073>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UN MODELE SPATIOTEMPOREL POUR ANALYSER LE COMMENCEMENT D'UNE EPIDEMIE, APPLICATION AUX ROUILLES DU BLE

A Spatiotemporal Model for Analyzing the Beginning of an Epidemic, Application to Wheat Rusts

Samuel **SOUBEYRAND***, Ivan **SACHE****
Michael **HÖHLE*****, Leonhard **HELD*****

* *Institut National de la Recherche Agronomique
Unité de Biométrie, Domaine Saint Paul, 84914 Avignon, France
Courriel : samuel.soubeyrand@avignon.inra.fr*

** *Institut National de la Recherche Agronomique
UMR Epidémiologie Végétale, 78850 Thiverval-Grignon, France
Courriel : ivan.sache@grignon.inra.fr*

*** *Universität Ludwig-Maximilians de Munich
Institut für Statistik, Ludwigstr. 33, 80333 München, Deutschland
Courriel : michael.hohle@stat.uni.muenchen.de
leonhard.held@stat.uni.muenchen.de*

SUMMARY

The beginning of an epidemic is often uncertain: it consists in few infections which lead to the emerging of the disease only if the local biological and environmental conditions are conducive. For understanding how an airborne plant disease such as the yellow rust of wheat emerges, epidemics in fields have been monitored in space and time, and a spatiotemporal model for analyzing data coming from these experiments have been developed. This document presents the ideas underlying the model and discusses its use for analyzing the beginning of an epidemic.

MOTIVATIONS-OBJECTIFS

Les rouilles du blé sont causées par des champignons qui forment des lésions sur les feuilles. Elles se propagent par l'intermédiaire de spores qui sont produites au sein des lésions et dispersées dans l'air principalement sous l'action du vent.

Le commencement d'une épidémie de rouille de blé est bien souvent balbutiant, incertain. En effet, il est constitué d'un nombre réduit d'attaques qui, selon les conditions biologiques et environnementales locales, causent ou non des infections ; de ce nombre réduit d'attaques dépend la réussite d'une émergence.

Des expériences consistant (i) à introduire une source de rouille brune ou jaune dans une parcelle de blé et (ii) à faire un suivi spatiotemporel de l'épidémie ont été effectuées par l'UMR Epidémiologie Végétale de l'INRA de Versailles-Grignon [1]. Les suivis consistaient à mesurer la sévérité de la maladie à des dates successives et sur un

ensemble d'unités d'observations (des groupes de feuilles) réparties dans l'espace 3D (plan horizontal x étages foliaires). Dans certaines expériences, un foyer de maladie s'est constitué autour de la source introduite puis la maladie s'est généralisée à la parcelle. Dans d'autres, un ou plusieurs foyers localisés ailleurs dans la parcelle ont été observés ; ces foyers, possiblement d'origine exogène, ont quelques fois dominé le foyer localisé autour de la source. De plus, il a été observé que certains foyers pouvaient régresser, voire disparaître.

Ces expériences ont permis d'observer des émergences, dues aux sources introduites ou à des sources exogènes, et des non-émergences. Afin de mener une analyse statistique des données issues de ces expériences et donc apprendre sur ce qui fait qu'une maladie émerge ou pas, nous avons construit un modèle spatiotemporel [2] adapté au caractère balbutiant du commencement d'une épidémie. De plus, ce modèle est adapté à l'analyse statistique de données de sévérité alors que les modèles proposés dans [3,4,5] ne permettent d'analyser que des données d'occurrence.

L'objectif de ce document est de présenter les idées qui sous-tendent notre modèle spatiotemporel et de voir en quoi il est adapté pour analyser quantitativement le commencement d'épidémies de maladies des plantes à dissémination aérienne.

METHODOLOGIE-RESULTATS

Le modèle spatiotemporel que nous avons développé décrit la variation dans le temps et l'espace de la sévérité qui est définie, pour une unité d'observation, comme la proportion de la surface foliaire sporulante, i.e. infectieuse. La sévérité, qui est une mesure quantitative de la maladie, est une variable particulièrement pertinente dans le cas des rouilles parce que c'est une bonne mesure du pouvoir infectieux d'une unité infectée.

Cependant, comme expliqué dans l'introduction, une épidémie a un aspect binaire : une unité d'observation est infectée ou ne l'est pas, et la mesure quantitative de la maladie qu'est la sévérité a une réalité seulement si l'unité est infectée. Notre modèle intègre les aspects binaire et quantitatif de l'épidémie : nous modélisons de manière jointe l'occurrence et la sévérité de la maladie. Dans ce document, l'occurrence est une variable qui vaut zéro si l'unité considérée n'a jamais été infectée et un sinon. Ainsi, inspiré par un modèle décrivant les événements pluvieux dans l'espace et le temps [6], nous modélisons à un premier niveau l'occurrence de la maladie, et à un deuxième niveau la sévérité de la maladie sachant l'état de l'occurrence.

Afin de présenter les idées qui sous-tendent le modèle, introduisons quelques notations. Soit $t = 1, \dots, T$ les temps d'observation. Soit $i = 1, \dots, n$ les unités d'observation sur lesquelles on mesure la sévérité de la maladie au cours du temps. Soit Y_{it} la sévérité de la maladie qui est observée sur l'unité i au temps t . Notons $Y_t = \{Y_{it} : i = 1, \dots, n\}$ l'ensemble des sévérités observées au temps t . Soit I_{it} la variable occurrence pour l'unité i au temps t . Dans les expériences mentionnées plus haut, c'est la sévérité qui est mesurée ; cependant, au vu de la définition de l'occurrence, quand on connaît la sévérité on connaît l'occurrence et donc, dans les faits, les valeurs des deux variables sont connues.

Dans notre modèle, la paire (I_{it}, Y_{it}) est modélisée (i) conditionnellement à l'état précédent $(I_{i,t-1}, Y_{i,t-1})$ afin de prendre en compte la situation de départ, mais également (ii) conditionnellement aux sévérités observées Y_{t-1} au temps précédent afin de prendre en compte l'effet des unités infectieuses sur l'unité i . La distribution jointe conditionnelle de (I_{it}, Y_{it}) est supposée vérifier

$$\Pr\left\{(I_{it}, Y_{it}) \mid (I_{i,t-1}, Y_{i,t-1}), Y_{t-1}\right\} = \Pr(Y_{it} \mid I_{it}, Y_{t-1}) \Pr(I_{it} \mid I_{i,t-1}, Y_{t-1}),$$

où $\Pr(I_{it} \mid I_{i,t-1}, Y_{t-1})$ représente le modèle pour l'occurrence et $\Pr(Y_{it} \mid I_{it}, Y_{t-1})$ le modèle pour la sévérité. Nous ne détaillerons pas ces modèles dans ce papier, nous

fournirons seulement des éléments utilisés pour les construire. Le détail se trouve dans [2].

Dans les modèles pour l'occurrence et la sévérité, le risque d'infection que font peser les unités infectieuses sur l'unité i est quantifié par un potentiel infectieux $W_{it} = \sum_{j=1}^n Y_{j,t-1} p_t(j,i)$, où $p_t(j,i)$ est la probabilité pour une spore produite et libérée par l'unité j d'être déposée sur l'unité i . Notons que $p_t(j,i)$ est supposée décroître avec la distance entre j et i . Dans l'expression du potentiel infectieux W_{it} inspirée par [7], on voit que chaque unité j contribue à W_{it} proportionnellement à la sévérité mesurée sur j au temps précédent et que cette contribution est modulée par la quantité $p_t(j,i)$ qui est d'autant plus grande que j est proche de i . La fonction $p_t(.,.)$, appelée fonction de dispersion, a été construite dans [2] comme une combinaison entre des fonctions de dispersion horizontale et verticale pour lesquelles différentes formes paramétriques ont été proposées.

Dans les modèles pour l'occurrence et la sévérité, des coefficients variant au cours du temps sont introduits afin de refléter l'effet de l'évolution des conditions environnementales sur la production de spores et l'efficacité d'infection notamment.

Modèle pour la variable occurrence I_{it} . Sachant $I_{i,t-1} = 1$, par définition $I_{it} = 1$. Sachant $I_{i,t-1} = 0$ et sachant les sévérités Y_{t-1} au temps $t-1$, I_{it} est supposé être distribuée selon une loi de Bernoulli : I_{it} reste à zéro avec une probabilité dépendant du potentiel infectieux W_{it} et d'un coefficient, noté α_t , variant dans le temps, et passe à un sinon. Plus les conditions sont favorables à la production de spores, à leur dispersion, et à l'efficacité d'infection, plus α_t est grand.

Modèle pour la variable sévérité Y_{it} . Sachant $I_{it} = 0$, par définition $Y_{it} = 0$. Sachant $I_{it} = 1$ et sachant les sévérités Y_{t-1} au temps $t-1$, Y_{it} est supposé être distribuée selon un mélange entre une loi de Dirac à zéro et une loi beta. La Dirac à zéro permet de modéliser les cas où la sévérité retourne à zéro après une période de stricte positivité. La loi beta est caractérisée par sa valeur moyenne et son paramètre d'échelle [8] qui sont supposés être fonction de la sévérité $Y_{i,t-1}$ sur i au temps précédent, du potentiel infectieux W_{it} et de coefficients variant dans le temps.

Le modèle spatiotemporel décrit ci-dessus a été utilisé pour analyser des données issues d'un suivi d'une épidémie de rouille jaune dans une parcelle de blé [9]. L'expérience a été conduite en 1996 sur le centre INRA de Versailles. Les données issues du suivi sont représentées sur la figure 1 dont la légende donne un bref descriptif de l'expérience. A partir de ces données, les paramètres du modèle ont été estimés par maximum de vraisemblance. La figure 2 montre l'évolution dans le temps des estimations du paramètre α_t qui intervient dans le modèle pour la variable occurrence. Sa faible valeur au temps 2 correspond à la quasi-stagnation de la maladie entre les temps 1 et 2. Sa forte valeur au temps 3 correspond à l'explosion de l'épidémie entre les temps 3 et 4. Et ainsi de suite.

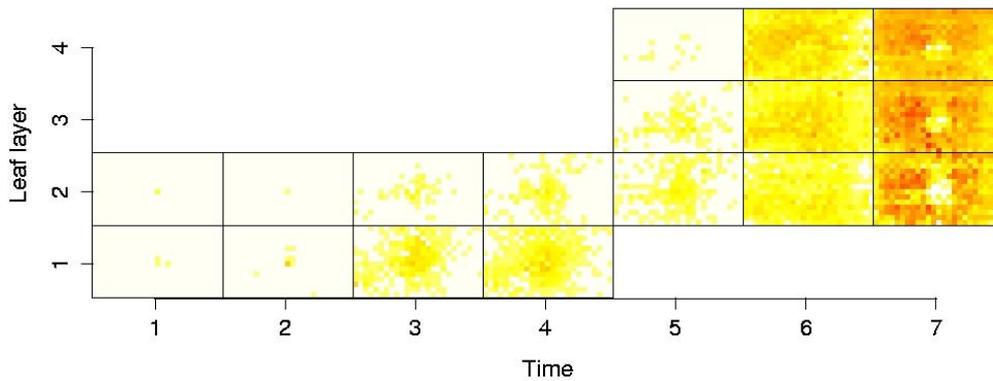


Fig 1. Evolution dans le temps et l'espace de la racine cubique de la sévérité mesurée lors du suivi spatio-temporel d'une épidémie de rouille jaune. La racine cubique est utilisée pour permettre de visualiser la quantité de maladie aux premières dates. L'épidémie a été initiée par le dépôt d'une source de rouille jaune au centre d'une parcelle de blé de 25x14m². La sévérité a été mesurée 46, 53, 61, 68, 76, 82, 89 jours après la mise en place de la source de rouille ; ces dates correspondent aux temps 1, 2, ..., 7. La sévérité a été mesurée sur des unités d'observation composées de 5 feuilles chacune et disposées sur un réseau tridimensionnel. Horizontalement le réseau était une grille rectangulaire avec une maille d'un mètre. Verticalement chaque unité d'observation était repérée selon son étage foliaire (1 (le plus bas), 2, 3 et 4). Entre les temps 1 et 4 il n'y a que deux étages foliaires ; entre les temps 4 et 5, les feuilles de l'étage 1 sont mortes et les feuilles des étages 3 et 4 ont émergé. Sur la figure, chaque rectangle fournit, pour une date donnée et un étage foliaire donné, les variations spatiales de la sévérité mesurée. Blanc : sévérité faible ; rouge : sévérité forte ; sévérité mesurée maximale : 0.62.

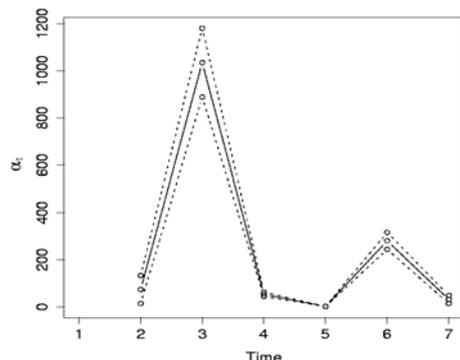


Fig. 2. Evolution dans le temps du paramètre α_t . Les points liés par les traits pleins sont les estimations de α_t . Entre les enveloppes constituées en pointillés : 4 écart-types. Il n'y a pas d'estimation au temps 1 parce que le modèle est conditionnel à la date précédente et qu'on ne connaît pas l'état du système au temps précédent le temps 1.

L'évolution temporelle d'un paramètre tel que α_t permet d'appréhender la dynamique globale de l'épidémie, de saisir comment celle-ci est modulée dans le temps. Mais afin d'appréhender de manière plus fine la dynamique épidémique, on peut s'intéresser à ce qui se passe dans l'espace à une date donnée. Par exemple, on remarque sur la figure 1 qu'aux premières dates la répartition spatiale de la maladie est éparse (présence de zones isolées de maladie). La figure 3 représente une épidémie simulée avec notre modèle spatiotemporel dont les paramètres ont été fixés aux valeurs estimées. On voit sur cette figure que notre modèle peut produire une répartition spatiale éparse de la maladie.

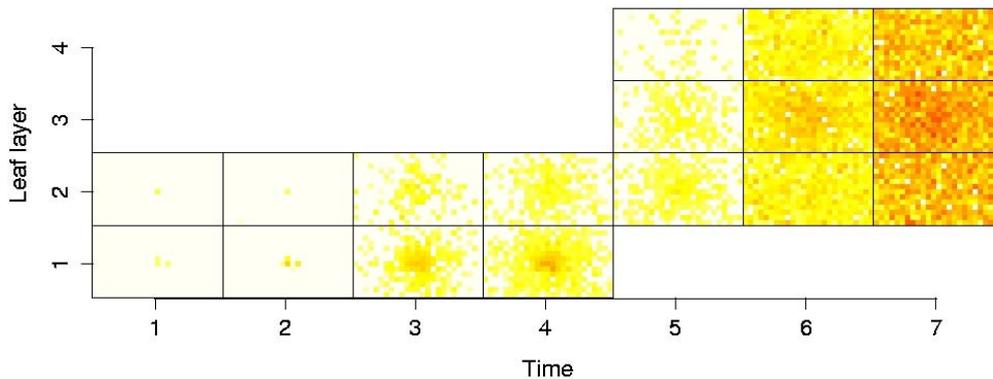


Fig. 3. Epidémie simulée selon le modèle spatiotemporel ajusté aux données. Les Conditions initiales correspondent aux sévérités observées au temps 1. La structure spatio-temporelle du couvert végétal a été laissée inchangée. L'échelle des couleurs est la même qu'à la figure 1.

DISCUSSION

Le modèle spatiotemporel présenté ci-dessus décrit de manière jointe l'occurrence et la sévérité d'une maladie à dissémination aérienne affectant des plantes. Dans ce modèle, le risque d'infection, dû aux unités infectieuses et pesant sur toutes les unités, est modélisé au travers d'un potentiel infectieux. Les accélérations, les stagnations, et les régressions globales de la maladie sont modélisées au travers de coefficients qui varient dans le temps.

L'épidémie qui a été analysée (brièvement ici mais plus profondément dans [2]) avec ce modèle présente une dynamique somme toute assez simple : grosso modo un foyer unique s'étend. On n'observe pas de foyer d'origine exogène ou encore pas de zone de la parcelle qui serait non propice à la maladie. Pourtant, dans l'ensemble des expériences effectuées par l'UMR Epidémiologie Végétale de l'INRA de Versailles-Grignon, de telles situations peuvent être observées. Le modèle pourrait être adapté à ces situations (i) en ajoutant au potentiel infectieux un terme reflétant le risque d'infection que font peser des sources exogènes non observées et (ii) en introduisant des coefficients qui varient dans l'espace tout comme on a introduit des coefficients qui varient dans le temps. On peut également penser à introduire des coefficients qui varient avec la génétique des plantes ou des souches de rouille. Ainsi le modèle gagnerait en flexibilité et pourrait mieux décrire les dynamiques épidémiques observées. Toutefois, il faut veiller à ce que l'augmentation du nombre de paramètres ne conduise pas à l'impossibilité d'estimer ces paramètres.

La modélisation jointe de l'occurrence et de la sévérité de la maladie est particulièrement intéressante car elle fait explicitement la différence entre un niveau de maladie nul et un niveau de maladie positif. Cette différence est fondamentale car (i) avec un niveau nul la maladie ne peut pas progresser, par contre (ii) avec un niveau positif, même faible, la maladie peut progresser. Parvenir à faire cette différence est particulièrement importante quand on s'intéresse au commencement d'une épidémie car substituer des plantes malades même faiblement malades à la place de plantes non malades peut résulter sur une dynamique épidémique profondément différente et certainement accélérée.

PERSPECTIVES-MESSAGE

Le modèle que nous avons présenté a été développé pour analyser des épidémies dans des parcelles. Ce travail peut être vu comme un banc d'essais à une échelle où le suivi d'une épidémie ne pose pas trop de problèmes. Toutefois, l'utilisation de notre modèle pourrait être étendue à une échelle spatiale plus grande afin d'étudier une émergence dans une région agricole par exemple. Dans ce cas là, la variable mesurée

pourrait être la sévérité de la maladie dans une parcelle agricole. Bien entendu il n'est pas envisageable de mesurer la sévérité dans toutes les parcelles d'une région, mais à l'échelle que nous avons étudiée dans ce document, le sondage n'est pas non plus exhaustif. Une difficulté plus sérieuse est la prise en compte de l'hétérogénéité des résistances des plantes et des virulences des agents pathogènes. Le contournement de cette difficulté devrait susciter d'intéressantes recherches.

BIBLIOGRAPHIE

1. Sache I., 2002. L'épidémiologie végétale : un avenir incertain pour une discipline ancienne ? Diplôme d'habilitation à diriger des recherches, Université Paris-Sud
2. Soubeyrand S., Sache I., Höhle M., Held L., 2006. Modelling the spread in space and time of an airborne plant disease. Research report, Unité Biométrie, Institut National de la Recherche Agronomique, Avignon
3. Chadœuf J., Nandris D., Geiger J.P., Nicole M., Pierrat J.C., 1992. Modélisation spatio-temporelle d'une épidémie par un processus de Gibbs : estimation et tests. *Biometrics* 48, 1165-1175
4. Gibson G.J., 1997. Investigating mechanisms of spatio-temporal epidemic spread using stochastic models. *Phytopathology* 87, 139-146
5. Gibson G.J., 1997. Markov chain Monte-Carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. *Journal of the Royal Statistical Society C* 46, 215-233
6. Chandler R.E., Yang C., Bellone E., Isham V., Northrop P., 2006. Space-time modelling of rainfall for continuous simulation. In Finkenstadt B., Held L., Isham V. (Eds), *Statistical Methods for Spatio-Temporal Systems*. London: Chapman & Hall
7. Mollison D., 1977. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society B* 39, 283-326
8. Stasinopoulos D.M., Rigby R.A., Akantziliotou C., 2006. Instructions on how to use the GAMLSS package in R. Technical report, STORM Research Centre, London Metropolitan University, London
9. Schermesser N., 1996. Analyse spatio-temporelle d'épidémies de rouille jaune du blé, causée par *Puccinia striiformis* west. Mémoire de DAA, Ecole Nationale Supérieure Agronomique de Rennes