Epidemic models for plants infection under mixed effects of temperature and wetness

Adil Soufi[1], Mustapha Ait Rami[2], Mustapha El Jarroudi[1], and Moussa El Jarroudi[3]

[1] University of Abdelmalek Essaadi, FST Tangier
Department of Mathematics, LMA,
adil.soufi@gmail.com, eljarroudi@hotmail.com
[2] ENSIT Engineering School Tangier 90000, Morocco
Laboratory Systems, Control & Decision
aitrami@ensi.ma
[3] University of Liège
Department of Environmental Sciences and Management
meljarroudi@uliege.be

Abstract. This paper deals with modeling and fitting for epedimic models and their applications to the field of plants disease. For this purpose, two models are proposed that are expressed as a blend of two functions which reflect the effect of the temperature and the wetness. In addition, we provide an original method to fit the proposed models by employing simple techniques that can constitute an easy-to-use tool for simulation, prediction and/or control. Moreover, the method accuracy and efficiency are evaluated for some reported works in the literature. Computational results are provided to show the validity and effectiveness of the proposed epidemic models for some plant infections.

keywords Plant Infections, Epidemic Models, Fitting, Nonlinear Optimization

1 Introduction

The conception of epedimic models for plant infections is one of the most important ingredients for disease forecasting and control [6, 1]. The infection evolutions can be evaluated via fundamental models that should be designed in order to reflet the interaction between some environmental variables captured in laboratory, greenhouse, field experiments or by simulation [6, 7, 10, 9]. As for instance, in the fields of crop physiology and agricultural meteorology: the temperature, the surface wetness, the humidity and the rainfall, are the commonly studied environmental variables [14, 12, 13].

In this paper, our contribution is twofold:
(a) Developing suitable infection models for plant pathogens.
(b) Proposing an original methodology for fitting.

Specefically, we provide two models that can be succesfully used for plant infections with mixed effects of the temperature and the wetness. Such models have an appealing advantage compared with other published models [2, 5, 4, 8]. Their whole parameters have an intrinsic biological meaning that describes the scale, the shape, and the location of the disease response.

A second advantage of the proposed approach is that unlikely for many other epedimic models reported in the literature for which the location parameters are predetermined or empirically fixed

or even absent in the model,(see for instance, [3, 2, 8]); is that, instead, these intrinsically biological parameters are treated as a core to be a priori designed.

Moreover, based on the introduced methodology, the proposed models were implemented and successfully compared to some well-known models in the literature [5, 4, 8]. Our original least-square procedure with bound constraints presented herein, yielded quite good and better fits compared to three kinds of plants with different fungal infections based upon experimental data sets. In all the cases, our models were very competitive and outperform the well-known reported models. As a result, we have shed a new light on an easy, accurate and efficient way for modelling and fitting the data for various fungal infections.

2 Modeling Plant Diseases

In this part, we perform a mixed effects modeling for plant disease through the relative infection response with regard to the effect of some environmental variables represented by the temperature and the wetness duration.

The first proposed model is defined as follows

$$y(t,w) \;=\; f_1(t,w) + \epsilon \tag{1}$$

where the function $f_1(t,w) :=$

$$\begin{cases} \alpha(t-t_I)^{\beta}(t_C-t)^{\gamma}[1-exp(-\kappa(w-w_I)^{\lambda}))] \text{ if } (t,w) \in [t_I\ t_C] \times [w_I\ +\infty[\\ 0 \quad \text{otherwise} \end{cases} \tag{2}$$

The second model can be viewed as a variant of the previous model (1) which is defined by

$$y(t,w) \;=\; f_2(t,w) + \epsilon \tag{3}$$

where the function $f_2(t,w) :=$

$$\begin{cases} \alpha(t-t_I)^{\beta}(t_C-t)^{\gamma}[1-exp(-\kappa(w-w_I)]^{\lambda} \text{ if } (t,w) \in [t_I\ t_C] \times [w_I\ +\infty[\\ 0 \quad \text{otherwise} \end{cases} \tag{4}$$

In fact one can see that the S-shaped function $1-exp(-\kappa(w-w_{min}))^{\lambda}$ in the model (1) is replaced by another S-shaped function of the form $[1-exp(-\kappa(w-w_I))]^{\lambda}$ in the model (3).

The involved entities of the proposed modeling are

- y is the mesure of the relative infection on a scale from 0 to 1 (or score response tacking values in the interval [0 1]).
- t is the temperature (in Celsius).
- w is wetness duration (in hours).
- ϵ is the model residual or perturbation error represented by an unknown random normal variable.

The parameters of the models (1) and (3) are defined by

- β, γ and λ are the shape paramerters.
- κ is a speed parameter.
- α is a scale parameter depending on the shape and location parameters.
- t_I, t_C and w_I are the location paramerters such that

· t_I : minimum temperature required for infection (in Celsius)

$$t_I := \min\{t \mid y \neq 0\}. \tag{5}$$

· t_C : maximum temperature required for cure (in Celsius)

$$t_C := \max\{t \mid y \neq 0\} \tag{6}$$

· w_I : minimum wetness duration for infection (in hours)

$$w_I := \min\{w \mid y \neq 0\}. \tag{7}$$

In the sequel, we shall show the features of the proposed models (1) and (3).

Remark 1. Note that by definition, the location parameters intrinsically satisfy the fact that for all (t, w) we have $y(t_I, w) = y(t_c, w) = y(t, w_I) = 0$.

Now, in order to fully caracterize our models the scale parameter α has to be designed. Specifically, this parameter should be selected such that the functions f_1 and f_2 take their values in the interval [0 1]. As a consequence, the proposed modeling is intrinsically based on the following result.

Proposition 1. *The function f_1 and f_2 take their values in the interval* [0 1] *if and only if* $\beta > 0$,$\gamma > 0$,$\lambda > 0$, $\kappa > 0$ *and the scale parameter α has the following form*

$$\alpha = \gamma^{-\beta}\beta^{-\gamma}(\beta + \gamma)^{\beta+\gamma}(t_C - t_I)^{-(\beta+\gamma)} \tag{8}$$

Proof 1 *The condition* $\beta > 0$,$\gamma > 0$,$\lambda > 0$, $\kappa > 0$ *results from the fact that the functions f_1 and f_2 do not go to infinity at the boundary of the domain* $[t_I\ t_C] \times [w_I$.

The rest of the proof can be shown based on the fact that

$$\sup\left(1 - exp(-\kappa(w - w_I)^{\lambda})\right) = \sup\left((1 - exp(-\kappa(w - w_I))^{\lambda}\right) = 1,$$

and

$$\max\left((t - t_I)^{\beta}(t_C - t)^{\gamma}\right) = \alpha^{-1}.$$

■

3 Models Identification

The identification procedure in the modeling of plants disease can be based on likelihood principle which under some statistically legitimate assumptions leads to fit a given model to the obseved data by applying the least square procedure.

Given the sample data of measured temperature $T := \{t_1, \ldots, t_n\}$ and the set of measured wetness duration $W := \{w_1, \ldots, w_n\}$ with their corresponding observed relative infection $Y := \{y_1, \ldots, y_n\}$. Then, in order to fit the model to the sample data Y by estimating the involved parameters, one has to look for $\beta, \gamma, \lambda, \kappa, t_I, t_C, w_I$ that are the optimal solution to the following nonlinear least square optimization (Sum of Squares Error (SSE) minimization)

$$\mathbf{SSE} := \min_{\beta,\gamma,\lambda,\kappa,t_I,t_C,w_I} \|y - \hat{y}\|^2 \tag{9}$$

In our case, $\hat{y} = f_1(t, w)$ or $\hat{y} = f_2(t, w)$ as defined before by equations (1)-(3) .

The above optimization problem do not incorporate any information on T and W. Henceforth, to tacke into account the features of Y and the model's conception, we introduce the following undelying constraints

$$\begin{cases} t_C^- \leq t_C \\ t_I \leq t_I^+ \\ w_I \leq w_I^+ \end{cases} \tag{10}$$

In this way, we exploit the structure of the optimization problem by incoporating the lower bound on t_C and the upper bounds respectively on t_I and w_I. In what follows, it is shown how to obtain them.

Proposition 2. *In order to fit the data sets (Y, T, W), the following bounds must hold*

$$\begin{cases} t_C^- := \max\limits_{y_i \neq 0}(t_i) \\ t_I^+ := \min\limits_{y_i \neq 0}(t_i) \\ w_I^+ := \min\limits_{y_i \neq 0}(w_i) \end{cases} \tag{11}$$

Proof 2 *The argument line is straightforward from the construction of the functions f_1 and f_2.* ∎

The observed null responses $y_i = 0$ are suitable for tightening the bounds given by the constraints (10). Specifically, one can exploit additional upper and lower bounds by considering the following bounds that result from the model's design.

Proposition 3. *Define the following subsets of $T := \{t_1, \ldots, t_n\}$ and $W := \{w_1, \ldots, w_n\}$ as*

$$\begin{cases} \mathcal{S}(t_C) := \{t_j \mid y_j = 0,\ t_j \geq t_C^-,\ w_j \geq w_I^+\} \\ \mathcal{S}(t_I) := \{t_j \mid y_j = 0,\ t_j \leq t_I^+,\ w_j \geq w_I^+\} \\ \mathcal{S}(w_I) := \{w_j \mid y_j = 0,\ t_I^+ \leq t_j \leq t_C^-,\ w_j \geq w_I^+\} \end{cases} \tag{12}$$

Assume that these subsets are not empty. Then, based upon the previous bounds t_C^-, t_I^+ and w_I^+ defined by (11), the following bounds hold true.

$$\begin{cases} t_C^- \leq t_C \leq t_C^+ \\ t_I^- \leq t_I \leq t_I^+ \\ w_I^- \leq w_I \leq w_I^+ \end{cases} \tag{13}$$

where

$$\begin{cases} t_C^+ := \min\limits_{t_j \in \mathcal{S}(t_C)}(t_j) \\ t_I^- := \max\limits_{t_j \in \mathcal{S}(t_I)}(t_j) \\ w_I^- := \max\limits_{t_j \in \mathcal{S}(w_I)}(w_j) \end{cases} \tag{14}$$

Moreover, in the case where the subsets $\mathcal{S}(t_C)$, $\mathcal{S}(t_I)$, $\mathcal{S}(w_I)$ are empty (or some of them), one can set t_C^+, t_I^- and w_I^- to some empirical or trivial meteorological extremum values. For instance, set $w_I^- = 0h$ and more or less $t_C^+ = 40C^\circ$, $t_I^- = 0C^\circ$.

Proof 3 *The argument line is a direct consequence of Proposition 2 with regard to the conception of the functions f_1 and f_2* ∎

Therefore, by appealing to Propositions 1.- 2.- 3., the constrained nonlinear least square optimization problem under consideration for fitting the proposed models to the sample data (Y, T, W) is suitably formulated as

$$\begin{aligned} &\min_{\beta,\gamma,\lambda,\kappa,t_I,t_C,w_I} \sum_{y_i \neq 0} (y_i - \hat{y}_i)^2 \\ &\textbf{subject to} \\ &\begin{cases} \beta > 0, \gamma > 0, \delta > 0, \kappa > 0 \\ t_C^- \leq t_C \leq t_C^+ \\ t_I^- \leq t_I \leq t_I^+ \\ w_I^- \leq w_I \leq w_I^+ \end{cases} \end{aligned} \tag{15}$$

This constrained optimization problem can usually be solved using nonlinear programming methods such as the classical Gauss-Newton with its many variants, Trust-Region, Interior Point method. Note that unlike linear least-squares method, nonlinear least square methods are not guaranteed to converge to the global minimizer due to the lack of convexity. But these methods can work well in practice, if the initial starting points are well-guessed or can be close to the optimal solution.

4 Models Validation and Numerical Results

In the sequel, we perform a comparison study with the most well-known nonlinear models related to fungal infection that are based upon the combined effects of wetness duration and temperature. These models can be summarized as follows

– **Beta model:**

$$f(t, w) = a(t - t_I)^b (t_C - t)^c w^d \tag{16}$$

– **Duthie's model:**

$$f(t, w) = \frac{e(h+1)h^{((h+1)^{-1}-1)} exp(g(h+1)^{-1}(t-f))}{1 + exp(g(t-f))} \tag{17}$$

– **Polynomial model:** There is many kinds such that the model $f(t, w)$ or the logit of y is a polynomial function $p(w, t)$ as

$$\log(y/(1-y)) = a_0 + w(a_1 + a_2 t + a_3 t^2 + a_4 t^3) \tag{18}$$

Appart from **SSE** index accurracy that yields close estimates to the real observed disease score y, the goodness of the fit in the simulation study greatly depends on the statistical measure of performance. Thus, it is essential in our validation that the selected model truly reflects the properties and the featurse of the infection evolution. This should be validated through common statistical goodness of fit tests such as the $\mathbf{R^2}$ and its adjusted value $\mathbf{R_a^2}$ which indicate the strength of the linear relationship between the predicted and observed values as they become closer to the value 1 (perfect fit).

For the fitting purpose, we have used interior point method to solve the optimization problem (15). This method is supported by Matlab via the function fmincon. We have chosen the initial

conditions t_I^0, t_C^0 and w_I^0 for the temperature and the wetness, close to the proposed bounds t_I^+, t_C^- and w_I^+. For the others initial parameters α^0,β^0,γ^0,λ^0 and κ^0 we have chosen values between 0 and 1.

The simulation results in Tables 1.and 2. show that the two proposed models are more effective than the other reported models [5, 4, 8]. It can be seen that our performance factors are the best since they provide a better fit and significantly outperform these alternative models.

Table 1: Different infection models based on published studies relating fungal infection to temperature and wetness duration are given with R-square (R^2), Adjusted R-square (R_a^2), the sum of squares due to error (SSE),Root mean squared error (RMSE), the signe $*$ means that not provided in the references.

Host and Pathogen	**Good of fitness**				
Immature Strawberry fruit by hytophthora cactorum	**Model**	R^2	R_a^2	SSE	$RMSE$
	Plynomial model	0.75	0.74	1.28	*
Medicinal Plant by Powdery Mildew	Beta model	0.81	*	*	*
Onion by Puccinia allii	Duthie Model	0.9501	*	*	*

Table 2: Estimation parameters from model 1 and 2 with their corresponding goodness of fit

Host and Pathogen	**Estimation parameters**									**Good of fitness**			
	Model	α	β	γ	κ	λ	t_I	t_C	w_I	R^2	R_a^2	SSE	$RMSE$
Immature Strawberry fruit by hytophthora cactorum	1	0.0646	1.0735	0.3244	0.1586	0.6774	5.3976	28.5732	2,8e-06	0.8145	0.8088	0.9807	0.1724
	2	0.0659	1.0676	0.3214	0.3187	0.7491	5.3878	28.5750	7.6e-06	0.8150	0.8094	0.9778	0.1721
Medicinal Plant by Powdery Mildew	1	2e-04	6.2009	2.0975	2.3e-06	0.4766	29.1952	37.5989	3.2591	0.9677	0.9662	0.1040	0.0687
	2	5.4e-07	10.0954	2.7882	1.3e-05	0.5013	27.0693	37.7911	2.8515	0.9664	0.9648	0.1083	0.0702
Onion by Puccinia allii	1	0.4805	0.1709	0.1370	2.1637	4.3576	2.9640	25.0000	0.6331	0.9549	0.9528	0.1728	0.0886
	2	0.2319	0.3289	0.2663	0.6727	3.8528	1.3060	25.0013	0.7237	0.9486	0.9462	0.1969	0.0946

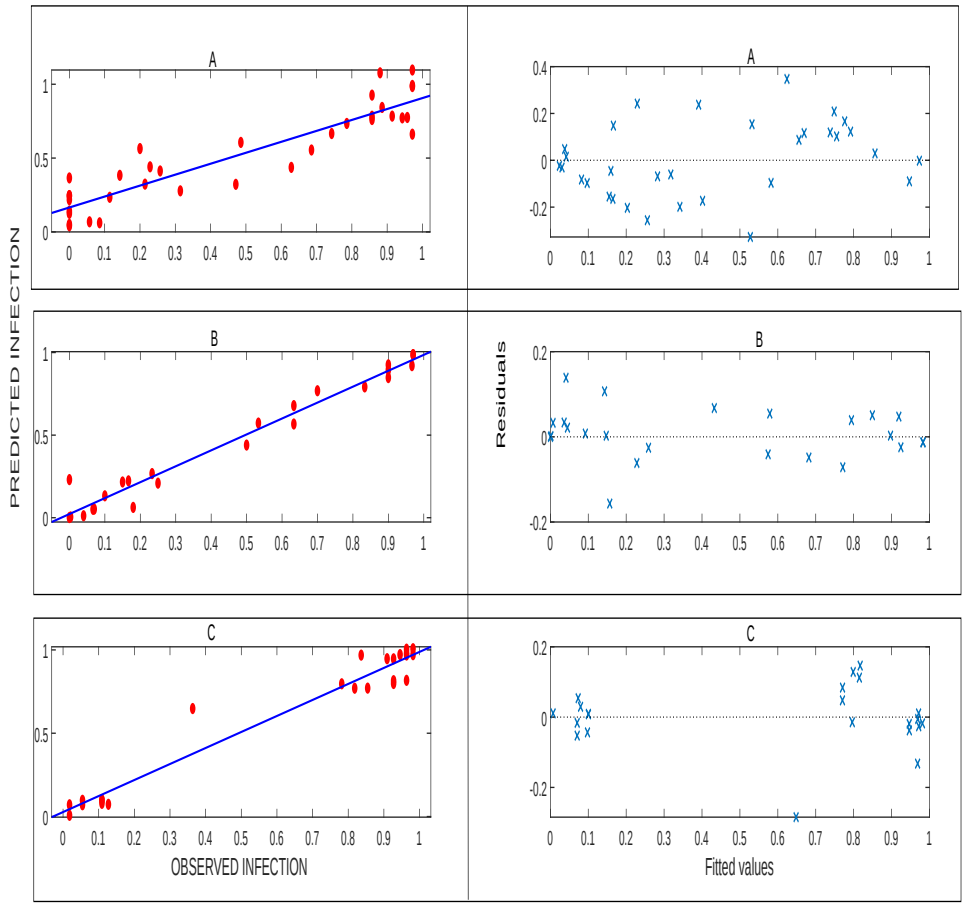

Fig. 1: Observed and predicted values of the relative disease using model 2 and Residuals versus the fitted values. A: Immature Strawberry fruit by hytophthora cactorum, B: Medicinal Plant by Powdery Mildew, C: Onion by Puccinia allii.

5 Conclusion

The method presented in this paper is original and can provide accurate modeling which is built upon simple techniques that may constitute an easy-to-use tool for simulation analysis, prediction and control for plant disease. In addition, the implementation procedure and the comparison study have shown the soundness of the proposed epedimic models.

References

1. Delignette-Muller ML. : Principle of Predictive Modeling. In F Toldra (ed.),Safety of Meat and Processed Meat, pp. 535-58. Springer-Verlag, New York (2009).
2. Duthie, J. A. : Models of the response of foliar parasites to the combined effects of temperature and duration of wetness. Phytopathology 87:1088-1095 (1997).
3. Erincik, O., Madden, LV, Ferree, DC, Ellis, MA. Temperature and wetness-duration requirements for grape leaf and cane infection by Phomopsis viticola. Plant Disease.87(7):832–40 (2003).
4. Furuya, H.,Takanashi, H., Fuji,S., Nagai, Y., and Naito, H.: Modeling Infection of !onion by Puccinia allii in Response to Temperature ans Leaf Wetness. The American Phytopathological, Vol. 99, No.8 (2009)
5. Grove GG, Madden LV, Ellis MA, Schmitthenner AF. : Influence of temperature and wetness duration on infection of immature strawberry fruit by Phytophthora cactorum.Phytopathology. 75(2):165–69 (1985).
6. Krause, R. A., and Massie, L. B. : Predictive systems : Modern approaches to disease control. Annu. Rev. Phytopathol. 13:31-47 (1975).
7. Madden, L. V., and Ellis, M. A. How to develop plant disease forecasters. Pages 191-208 in: Experimental Techniques in Plant Disease,(1988).
8. Sharma R.,V.D.Shivling, D. Kumar, A. K. Sharma : A Beta Regression Model for Himalayan Medicinal Plant Disease Prediction.Indian Journal of Science and Technology, Vol 7(6), 776–780 (2014).
9. Shrum, R., D., : Forecasting of epidemics. Pages 223-238 in: Plant Disease: an Advanced Treatise. J. G. Horsfall and E. B. Cowling, eds. Academic Press, New York (1978).
10. Waggoner, P., E. : Simulation of epidemics. Pages 291-312 in: Epidemics of Plant Diseases. J. Kranz, ed. Chapman and Hall, New York (1974).
11. Wilson, L., L., Madden, L., V., and Ellis, M., A. : Influence of temperature and wetness duration on infection of immature and mature strawberry fruit by Colletotrichum acutatum. Phytopathology 80: 111-116 (1990).
12. Yan, W., and Hunt, L., A. : An equation for modelling the temperature response of plants using only the cardinal temperatures. Ann. Bot. 84:607-614 (1999)..
13. Yin, X., Kropff, M., J., Mclaren, G., and Visperas, R., M. : A nonlinear model for crop development as a function of temperature. Agric. For. Meteorol. 77:1-16 (1995).
14. Wang, E., L., and Engel, T. : Simulation of phenological development of wheat crops. Agric. Sys. 58:1-24 (1998).