Model-based estimation of the dates and places of introduction of new WMV viral strains

L Roques, C Desbiez, K Berthier, S Soubeyrand, E Walker, E K Klein, J Garnier, B Moury, J Papaïx – MEDIA network

General context

Where and when alien organisms are successfully introduced are central questions for the study of biological invasions, to elucidate biotic and abiotic conditions favorable to the introduction, establishment and spread of invasive species:

- map the risk of invasion and design a more efficient surveillance,
- control of invaders by management strategies is easiest at early stages,
- precise knowledge of the dates and places of introduction is critical to estimate accurately the reproduction and dispersal parameters…

Biological invasions by alien organisms are often reported several years after the initial successful introduction event: **how to reconstruct the date and place of introduction?**

The Watermelon Mosaic Virus

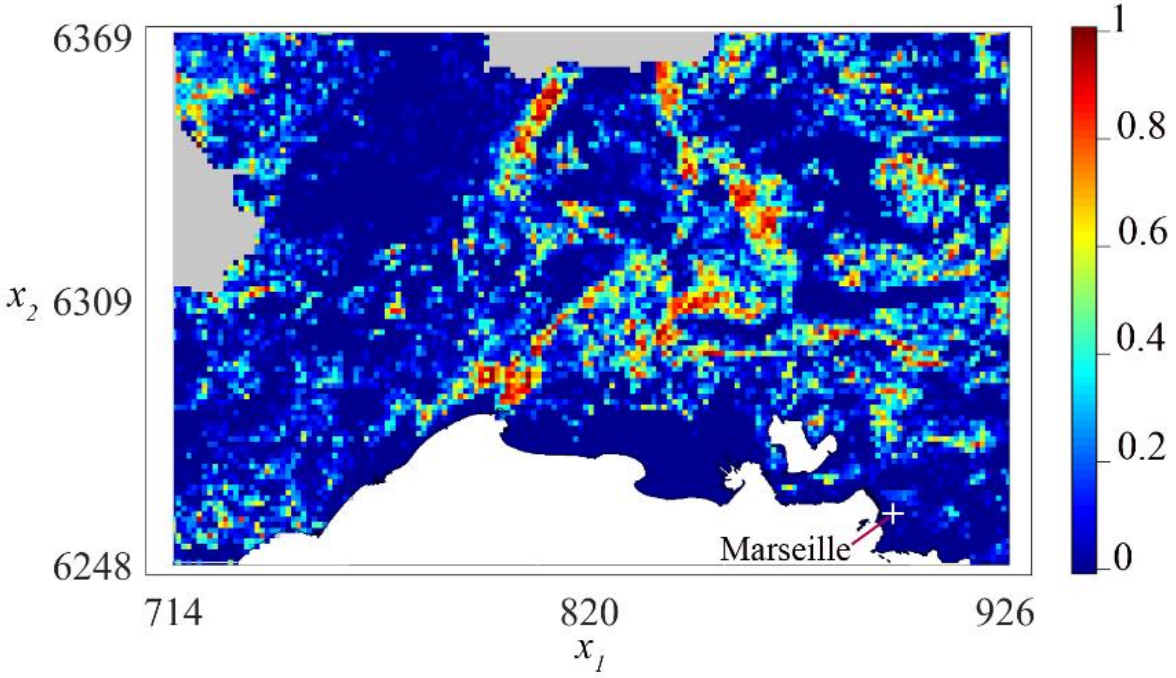
Before 1999

After 1999

Datasets

	2004	2005	2006	2007	2008
# observations	67	64	68	50	40
# infected samples	408	371	422	280	212
Classical strain	55%	45%	28%	17%	14%
Emerging strain 1	21%	23%	22%	37%	27%
Emerging strain 2	13%	18%	23%	21%	32%
Emerging strain 3	1%	4%	3%	5%	3%
Emerging strain 4	10%	10%	24%	20%	24%

Favorable habitat for WMV



Mechanistic model

The model describes the **spatio-temporal dynamics of the virus strains** (at position $\boldsymbol{x}$ and at time t of year n):

- $c^n(t, \boldsymbol{x})$ density of classical strain (CS)
- $e_k^n(t, \boldsymbol{x})$ densities of emerging strains (ES_k, indexed by $k = 1, \ldots, 4$).

The model is segmented into two stages:

1. **intra-annual stage**, May 1st ($t = 0$) to September 30th ($t_f = 153$ days), dispersal, growth and competition of the five virus strains during the summer epidemics on cucurbit crops.
2. **inter-annual stage**, discrete time, winter mortality of the different strains when no crops are present and the virus overwinters in weeds.

Mechanistic model

Dynamics of the classical strain before the first introduction events

$$\partial_t c^n = D\Delta c^n + r c^n (z(\boldsymbol{x}) - c^n)$$

- Δ: Laplace 2D diffusion operator
- D: diffusion parameter measuring the mobility of the viruses
- $r\ z(\boldsymbol{x})$: intrinsic growth rate (i.e., growth rate in the absence of competition). The carrying capacity at a position $\boldsymbol{x}$ is equal to $z(\boldsymbol{x})$, which means that the population densities are expressed in units of the maximum host population density.

Initialization: $c^{1980}(0, \boldsymbol{x}) = (1 - m_c)\ z(\boldsymbol{x})$, where m_c is the winter mortality rate of the CS (CS density is at the carrying capacity in 1979).

Mechanistic model

Introduction events

The ESs are introduced during years noted $n_k \geq 1981$, at the beginning of the intra-annual stage:

- $e_k^n = 0$ for $n < n_k$
- initial density of any ES, $e_k^{n_k}(0, x) = \frac{z(x)}{10} \exp\left(-\frac{\|x - X_k\|^2}{2\sigma^2}\right)$, where $\boldsymbol{X_k}$ is the location of introduction of the strain k.

Mechanistic model

Intra-annual dynamics

Intra-annual dynamics were described by a **neutral competition model with diffusion**, for $t = 0 \dots t_f$ and for all introduced emerging strains, i.e. all k such that $n \geq n_k$:

$$\begin{cases} \partial_t c^n(t,\boldsymbol{x}) = D\Delta c^n + r c^n \left(z(\boldsymbol{x}) - c^n - \sum_{i=1}^{4} e_i^n(t,\boldsymbol{x}) \right) \\ \partial_t e_k^n(t,\boldsymbol{x}) = D\Delta e_k^n + r e_k^n \left(z(\boldsymbol{x}) - c^n - \sum_{i=1}^{4} e_i^n(t,\boldsymbol{x}) \right) \end{cases},$$

- We assume reflecting boundary conditions
- Diffusion, competition and growth coefficients are common to all the strains during the intra-annual stage

Mechanistic model

Inter-annual dynamics

The population densities at time $t = 0$ of year n are connected with those of year $n - 1$, at time $t = t_f$, through the following formulas:

$$\begin{cases} c^n(0, \boldsymbol{x}) = (1 - m_c)c^{n-1}(t_f, \boldsymbol{x}) \text{ for } n \geq 1981 \\ e_k^n(0, \boldsymbol{x}) = (1 - m_e)e_k^{n-1}(t_f, \boldsymbol{x}) \text{ for } n \geq n_k + 1 \end{cases}$$

- m_c the winter mortality rates of the CS strain,
- m_e the winter mortality rates of the ESs strains (m_e is common to all of the ESs).

Probabilistic model for the observations

- C_i and $E_{k,i}$ for $k = 1, \dots, 4$ and $i = 1, \dots, I_n$: number of samples infected by the classical and emerging strains at each date of observation and location $(t_i, \boldsymbol{x}_i)$.
- $V_i = C_i + \sum_{k=1}^{4} E_{k,i}$: total number of infected samples observed at $(t_i, \boldsymbol{x}_i)$
- $\boldsymbol{p}_i = (p_i^c, p_i^{e_1}, p_i^{e_2}, p_i^{e_3}, p_i^{e_4})$: proportions of each strain in the population at $(t_i, \boldsymbol{x}_i)$,

$$p_i^c(\boldsymbol{\Theta}) = \frac{c^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta})}{c^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta}) + \sum_{i=1}^{4} e_i^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta})}, p_i^{e_k}(\boldsymbol{\Theta}) = \frac{e_k^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta})}{c^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta}) + \sum_{i=1}^{4} e_i^n(t_i, \boldsymbol{x}_i|\boldsymbol{\Theta})}$$

$$(C_i, E_{1,i}, E_{2,i}, E_{3,i}, E_{4,i}) \sim \mathcal{M}(V_i, \boldsymbol{p}_i)$$

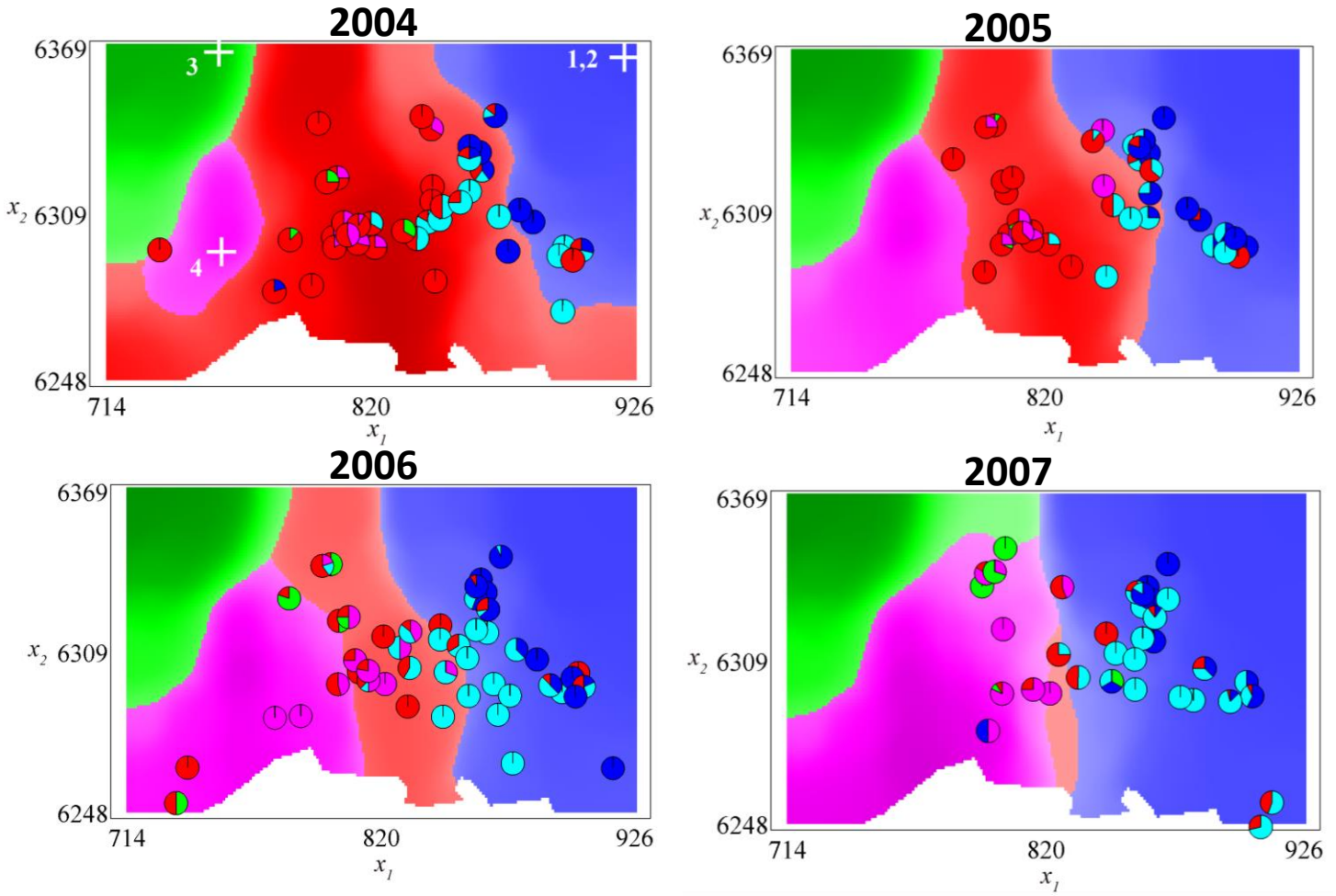
Statistical inference

- We used a simulated annealing algorithm to compute the maximum likelihood estimate.
- We ran 6 parallel sequences with cooling rates $\alpha \in \{0.995, 0.999, 0.9995\}$.
- Stopping criterion was that $\boldsymbol{\Theta}_j$ remained unchanged during 500 iterations.
- The computations took about 100 days (CPU time).

Results

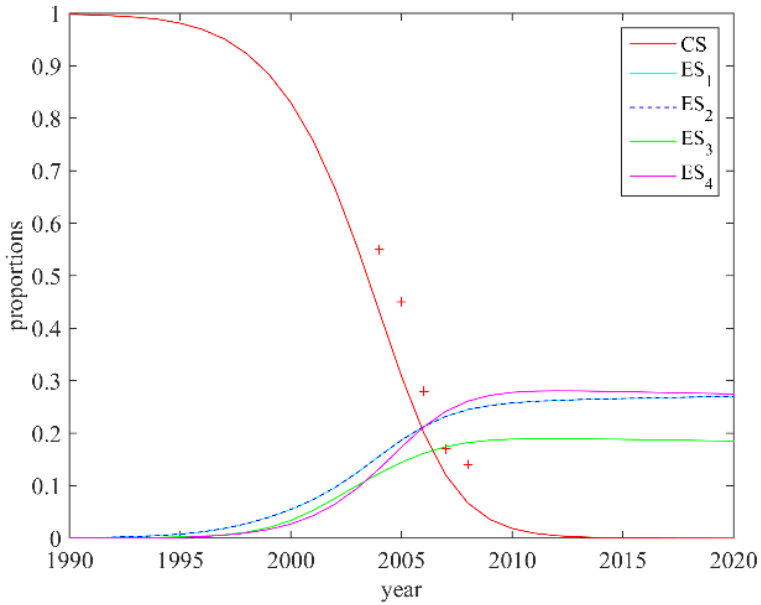
Biological parameter	D^*	r^*	m_c^*	m_e^*
Value	$0.44\ \mathrm{km}^2\ \mathrm{day}^{-1}$	$0.31\ \mathrm{day}^{-1}$	$0.5\ \mathrm{year}^{-1}$	$0\ \mathrm{year}^{-1}$
Date of introduction	n_1^* (ES1)	$n_{2\ (ES2)}^*$	$n_3^*\ (ES3)$	$n_4^*\ (ES4)$
Value	1990	1990	1990	1995
Site of introduction	$\boldsymbol{X}_1^*$	$\boldsymbol{X}_2^*$	$\boldsymbol{X}_3^*$	$\boldsymbol{X}_4^*$
Value (Lambert 93, km)	(926,6369)	(926,6369)	(758,6369)	(758,6294)

Results

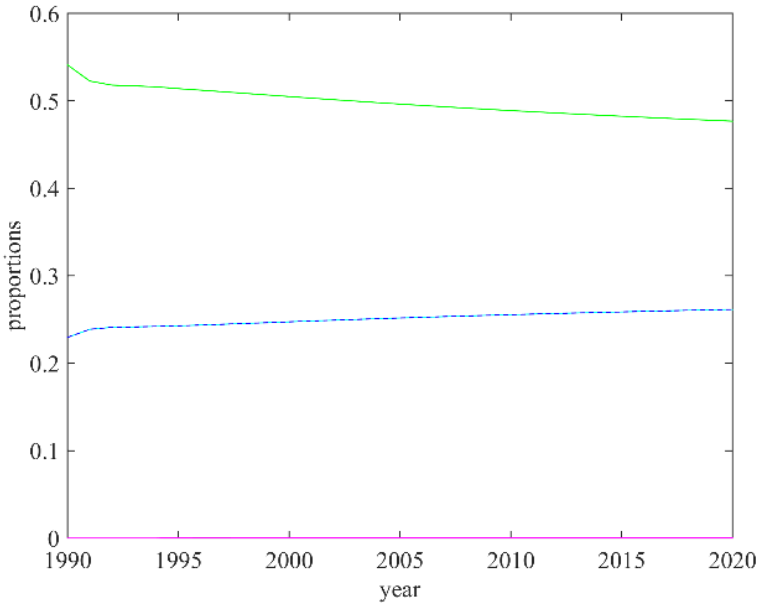


Results

Estimated average proportions of the classical and emerging strains



Dynamics of emerging strains in the absence of the classical strain



Discussion

We developed a reaction-diffusion model to describe the **spatial dynamics of invasion** of a resident population inhabiting a spatially structured environment by newly introduced variants and:

- estimated the **dates and places of successful introduction** of each emerging variant, as well as parameters related to **growth and dispersal**,
- **reconstructed the invasion** by the new variants from their introduction sites,
- established a **competitive advantage of the new variants** as compared to the resident population
- predicted the **fate of each variant**.

Perspectives:

- emerging strains differed only through their winter mortlity rate…
- better estimation of parameter uncertainty