Using survival analysis to assess grape downy mildew onset in Bordeaux vineyards

Mathilde CHEN (ACTA- INRA)

François BRUN (ACTA), Marc RAYNAL (IFV) & David Makowski (INRA)

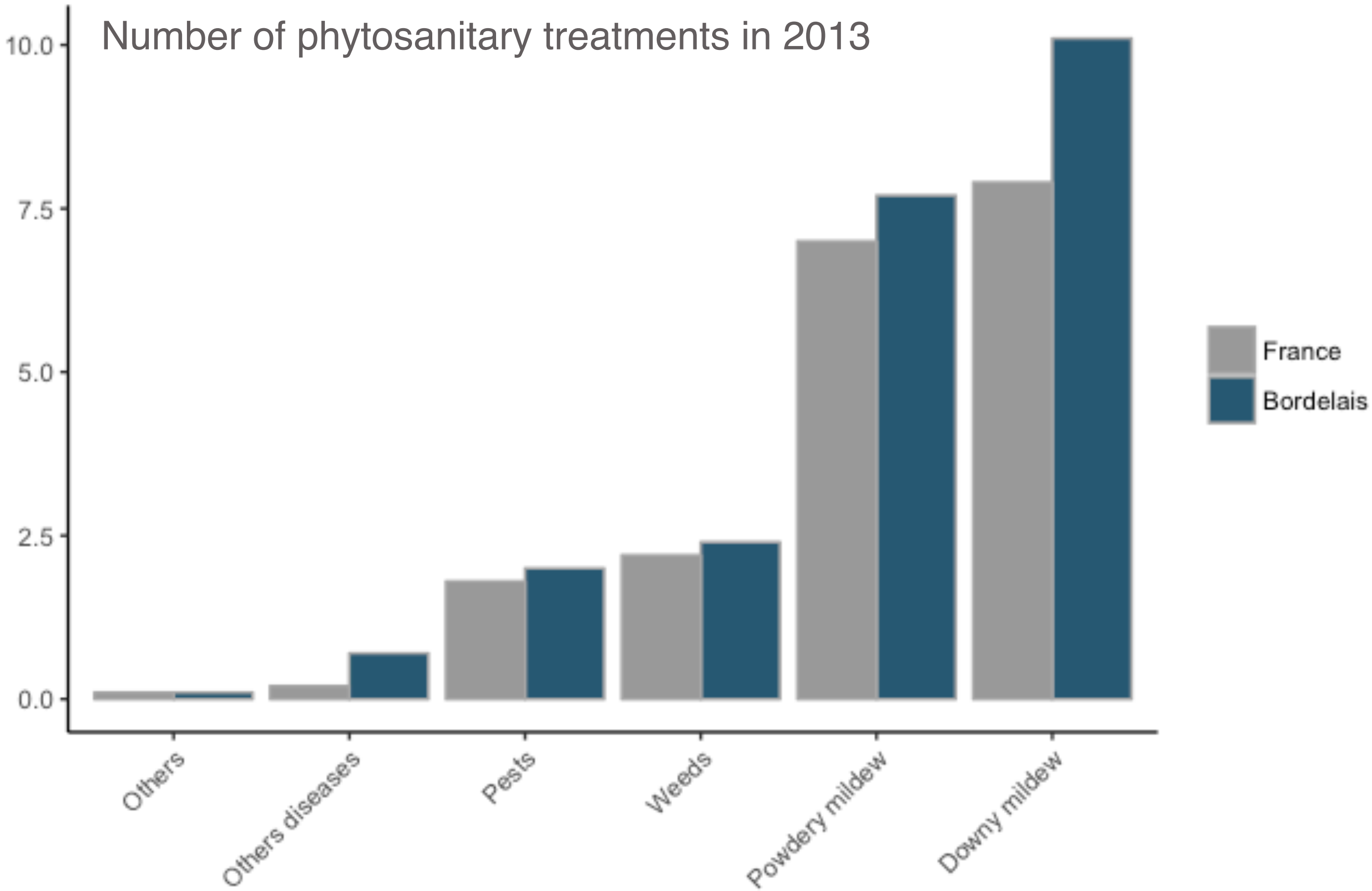
ACTA LES INSTITUTS TECHNIQUES AGRICOLES

INRA SCIENCE & IMPACT

IFV INSTITUT FRANÇAIS DE LA VIGNE ET DU VIN

ModStatSAP – Annual meeting
2018-03-19

Intensive control of vine's health



Source: Agreste – Enquête sur les pratiques culturales en viticulture 2013

Grape downy mildew control strategy

Widely recommended:

- a first fungicide application at first symptoms appearance
- regularly repeated sprayings until the end of the period of susceptibility

Grape downy mildew control strategy

Widely recommended:

- a first fungicide application at first symptoms appearance → impacts on the number of treatments
- regularly repeated sprayings until the end of the period of susceptibility

Source:Gessler et al., 2011; Mailly et al., 2016

Question

When do first symptoms occur on vines and on bunches in Bordelais vineyards ?

Data

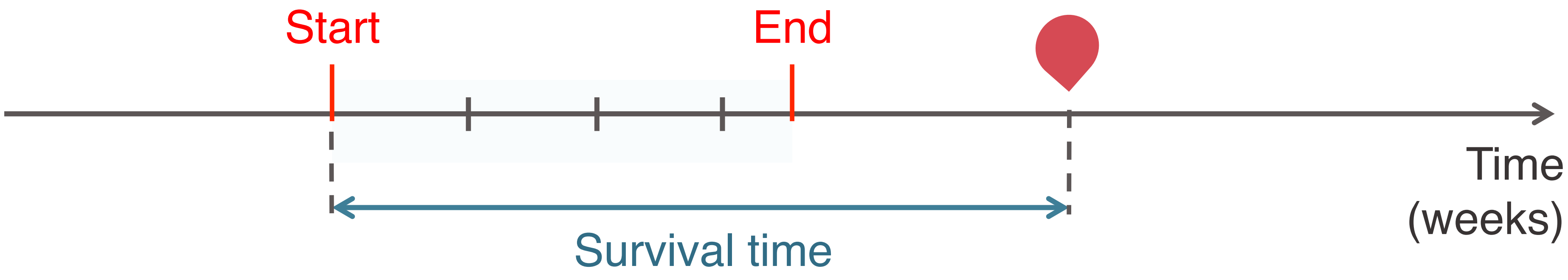
Incidence data

- Collected weekly
- Between 2010 and 2017
- In **non-treated plots**

No-censored data

Right-censored data

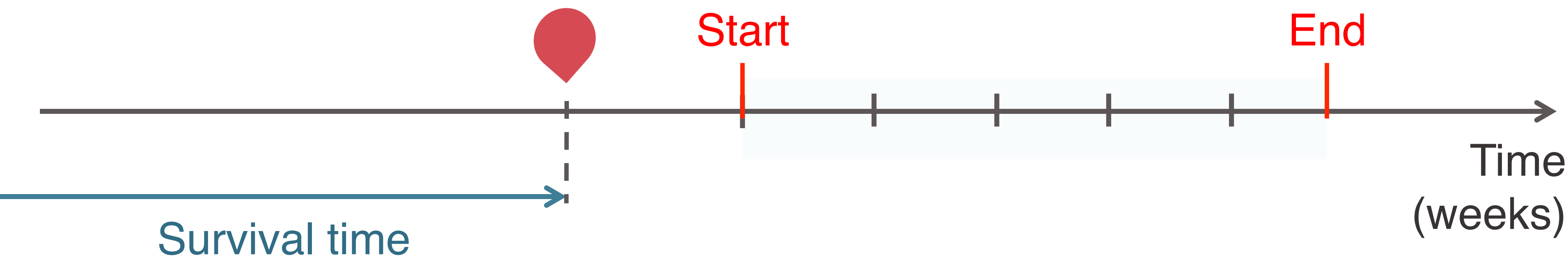
Survival time is longer than monitoring length



Example: a plot still symptomless at the end of the season

Left-censored data

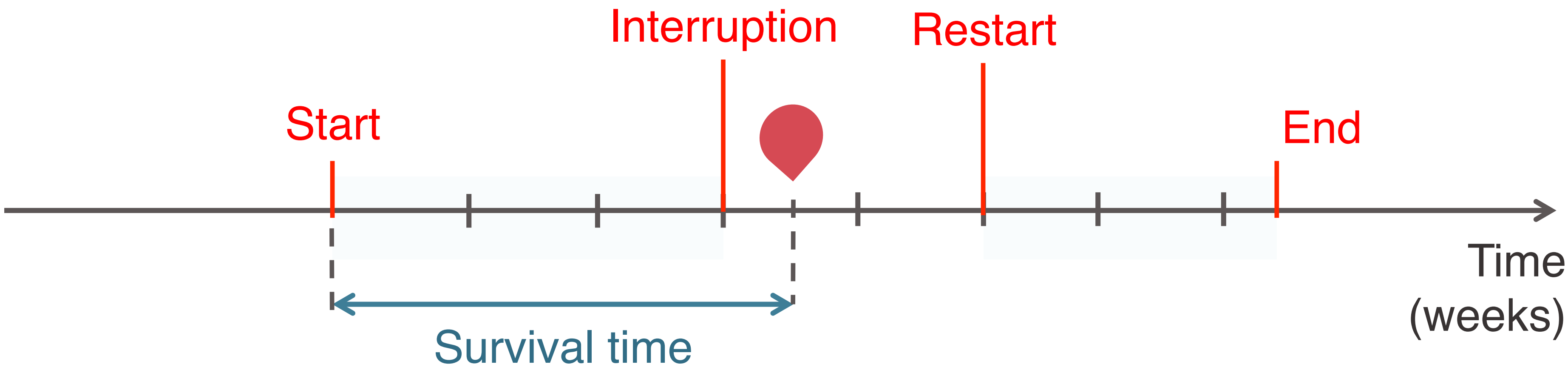
Event occured before monitoring start



Example: the plot is not symptomless at the first observation

Interval-censored data

Event occurs in a time interval



Example: the observer is sick during 3 weeks, the symptoms appeared when he comes back

Statistical analysis

Censored data represent 74% of the dataset

→ Using **survival analysis** to assess the expected time between the 01/01 and the first symptoms occurrence

Source: Scherm and Ojiambo (2004)

Definition of the survival and the hazard function

Survival function S(t)

S(t): proportion of symptomless plots

T: the time to GDM symptom appearance on 1% of vines or bunches

$$S(t) = P(T > t)$$

P(T > t): the probability that this time exceeds t in a plot

Definition of the survival and the hazard function

Hazard function $h(t)$

h(t): instantaneous rate of first symptom appearance at time t

$$h(t) = f(t) \,/\, S(t)$$

f(t): the probability density function

S(t): proportion of symptomless plots

Several estimation methods

Non-parametric estimation of S(t)

Kaplan-Meier estimator

$$S(t_i) = \left(\frac{n_1 - d_1}{n_1}\right)\left(\frac{n_2 - d_2}{n_2}\right)\cdots\left(\frac{n_i - d_i}{n_i}\right)$$

n_i : number of symptomless and non-censored plots just before the i^{th} observation
d_i : number of « infected » plots at the i^{th} observation

Generalization: Turnbull's estimator or NPMLE

Semi-parametric estimation of h(t)

Cox model

$$h(t, X, \beta) = h_0(t)\, e^{X'\beta}$$

$h_0(t)$: non-parametric baseline function

X: a set of covariate
β: a set of parameters

Model fitter with year effect as a covariate (X)
For bunches: symptom onset date on vine as a covariate

Parametric estimation of S(t)

Parametric models: log logistic example

$$ln(T) = \beta_0 + X'\beta + \sigma Z$$

β_0 and σ: parameters of the log logistic distribution

X: set of covariates
β: set of parameters

Z: baseline hazard function defined as a random variable

R packages for survival analysis

- survival

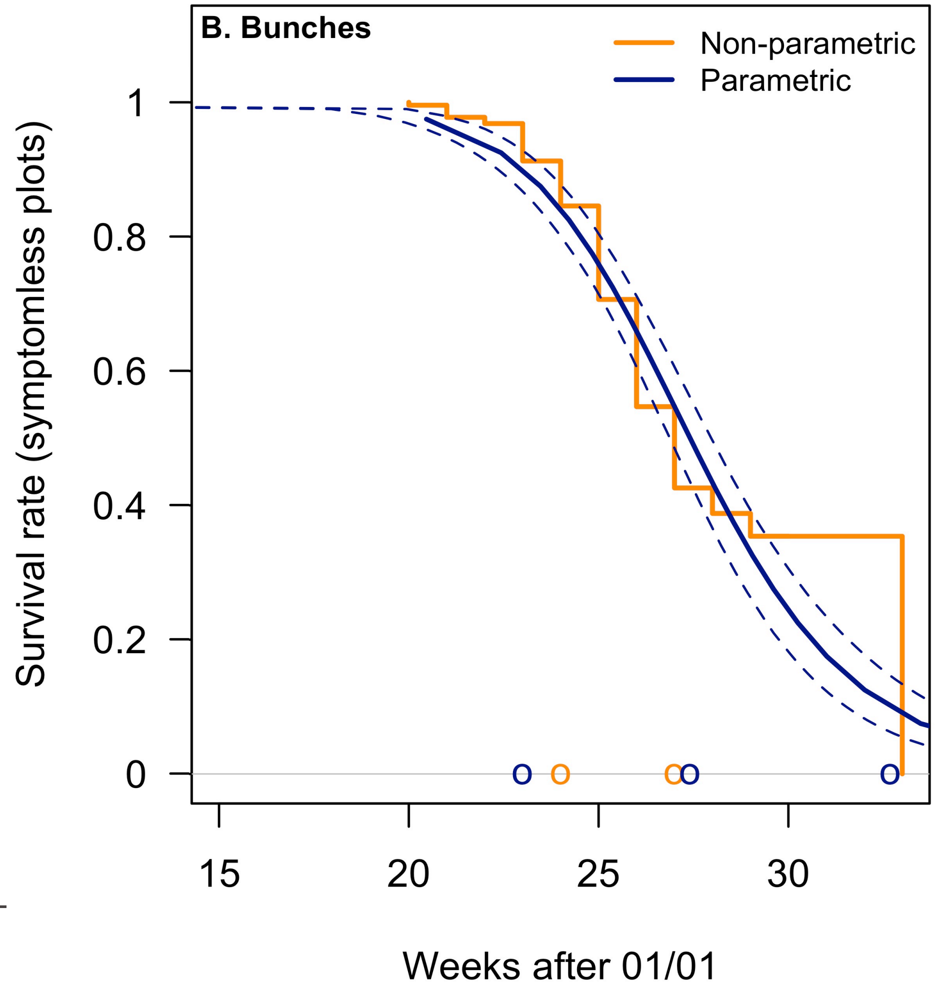
Therneau T. (2015)

- icenReg

Anderson-Bergman C. (2017)

→ Manage interval-censored data

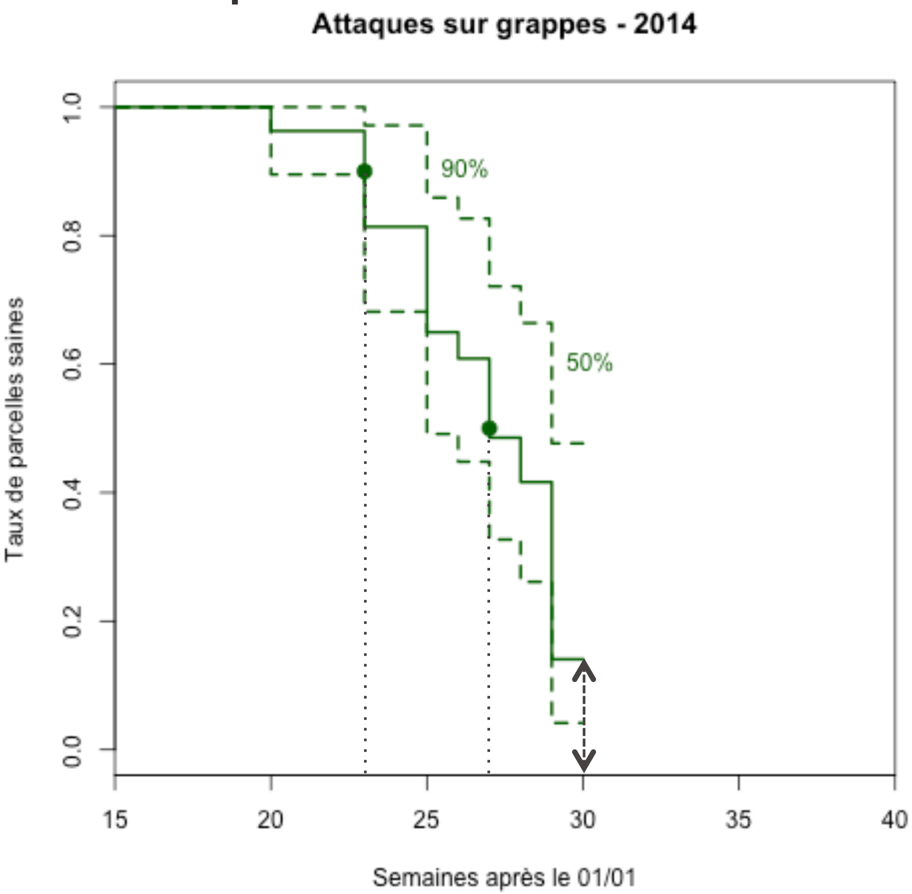
Results estimation of S(t) for bunches

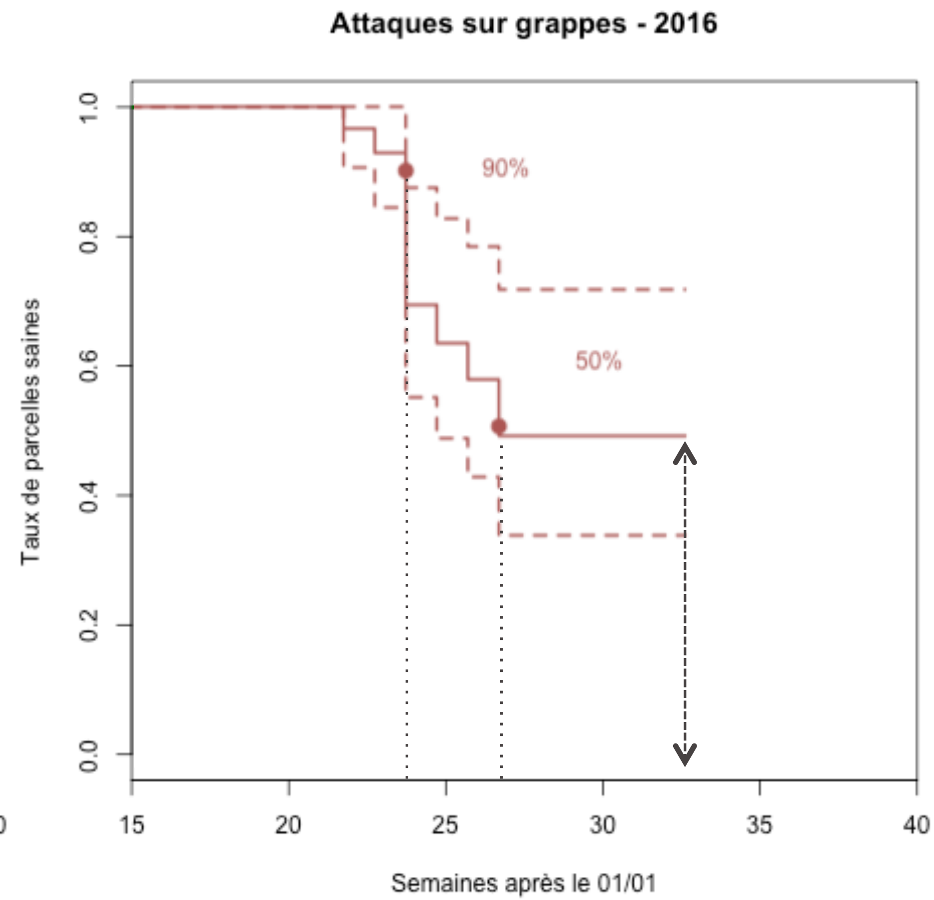


Source: Chen et al.

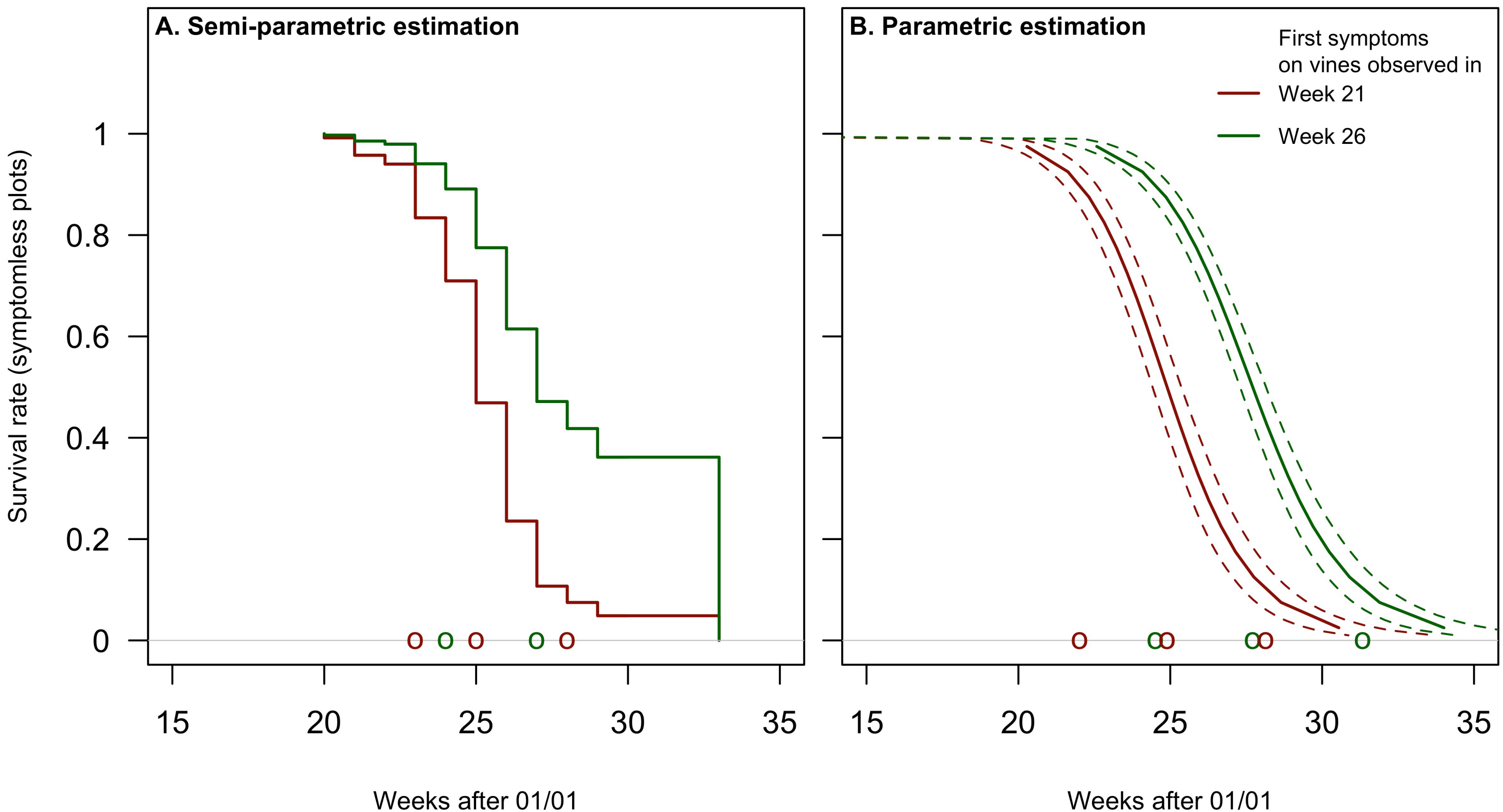
Variability of S(t) between years

Example for bunches





Relationship between GDM appearance on vines et bunches



Source: Chen et al.

Conclusions

- GDM onset may begin in May
- But may never occur

For several years, our statistical analysis revealed that the proportion of plots with no symptoms was high in early August on bunches

Conclusions

- Appearance rate is variable across years

Depending on the year, onset was recorded between mid-May late June for 50% of the plots

Thank you for your attention

Literature

Ambiaud, E., 2015. Enquête Pratiques culturales en viticulture 2013. Nombre de traitements phytosanitaires.

Anderson-Bergman, C., 2017a. **icenReg** : Regression Models for Interval Censored Data in *R*. Journal of Statistical Software 81. https://doi.org/10.18637/jss.v081.i12

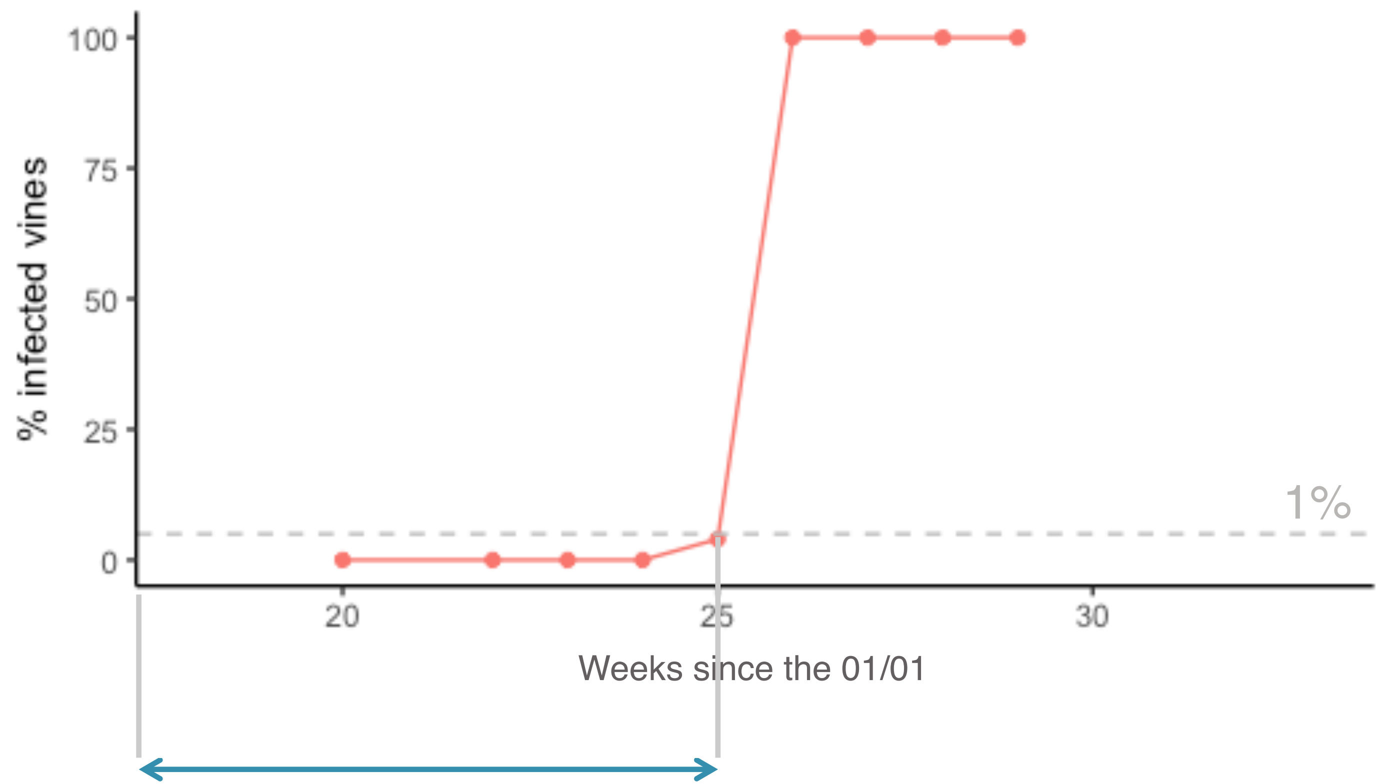
Gessler, C., Pertot, I., Perazzolli, M., 2011. Plasmopara viticola: a review of knowledge on downy mildew of grapevine and effective disease management. Phytopathologia Mediterranea 50, 3–44.

Mailly, F., Hossard, L., Barbier, J.-M., Thiollet-Scholtus, M., Gary, C., 2017. Quantifying the impact of crop protection practices on pesticide use in wine-growing systems. European Journal of Agronomy 84, 23–34. https://doi.org/10.1016/j.eja.2016.12.005

Scherm, H., Ojiambo, P.S., 2004. Applications of survival analysis in botanical epidemiology. Phytopathology 94, 1022–1026.

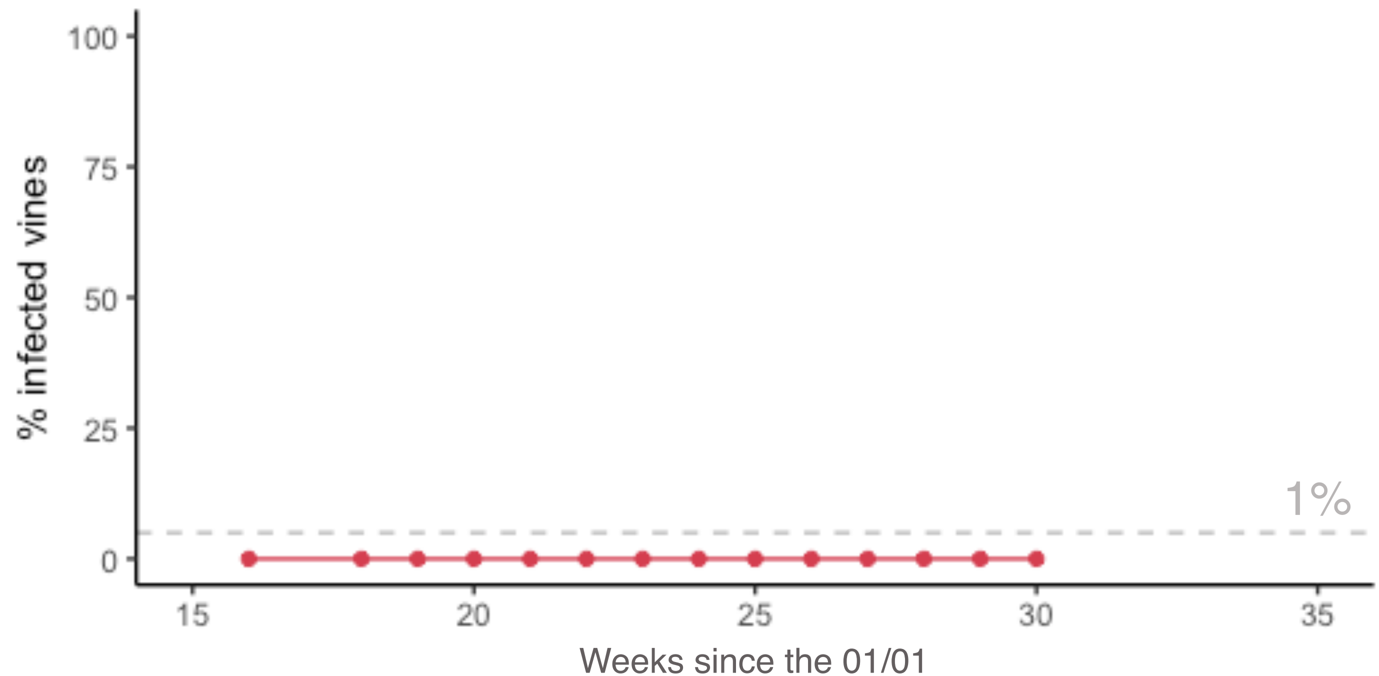
Therneau T (2015). *A Package for Survival Analysis in S*. version 2.38, https://CRAN.R-project.org/package=survival.

No-censored data



Time to GDM symptom appearance on 1% of vines or bunches (T)
= Survival time

Right-censored data



Left-censored data