Journal de la Société Française de Statistique
Vol. 157 No. 1 (2016)

Editorial for the special issue “Modelling and Inference for Infectious diseases”

Titre: Éditorial pour le numéro spécial "Modélisation et inférence pour les maladies infectieuses"

Sylvie Huet[1], Estelle Kuhn[1], Catherine Larédo[1] and Elisabeta Vergu[1]

Most of the main questions in the field of infectious diseases are related to the early detection of outbreaks, the spatial and temporal spread of infections and their control. Data are usually collected through direct surveillance of cases or through syndromic surveillance (health-related data available before infection confirmation). Most often, data are partially observed (including only a fraction of infected individuals, . . .), temporally or spatially aggregated (numbers of new cases over time for example) and heterogeneous (pathogen genomic data, indirect indicators, . . .). In this context, mathematical models coupled with statistical methods able to integrate various sources of information are useful tools for understanding and predicting spatio-temporal infectious disease dynamics, and for estimating key parameters in absence of complete data.

The five papers of this special issue are focused on stochastic models, statistical methods and algorithmic developments for the description and the analysis of mechanisms and data related to infectious diseases. The considered issues and methods developed may be of interest in human or veterinary epidemiology as well as plant disease epidemiology.

This special issue starts with an introduction by Britton and Giardina to the statistical inference for infectious diseases. The authors present stochastic epidemic models for the spread of infectious diseases, starting with the simple SIR model. Both individual heterogeneities and heterogeneous mixing are discussed. The authors also provide an overview of statistical methods for estimating model parameters assuming different types of data. A particular focus is made on estimation methods for partially observed data. Moreover, the authors show how statistical models can be used for early outbreak detection in disease surveillance and forecasting.

Guy *et al.* propose new inference methods based on diffusion processes approximating epidemic dynamics. Multidimensional diffusion processes with small diffusion coefficient are obtained: each component represents a health status of the epidemic dynamics. They first study the case where all the components are observed at discrete times and consider then the case where some components are not observed. Estimators of the parameters of the diffusion process are proposed. Their theoretical properties are established considering different patterns of data and various asymptotic frameworks. Numerical experiments on SIR and SIRS models using simulated and real data are used to assess the performances of the estimators.

Jacob and Pénisson use a multi-type branching process model for analysing the dynamics of

---

[1] MaIAGE, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France.
E-mail: `sylvie.huet@jouy.inra.fr`

the Bovine Spongiform Encephalopathy epidemic in Great-Britain. For both the growth phase and the decay phase of the epidemic, the authors propose estimators of the model parameters allowing to predict the number of new cases, the time to epidemic extinction and the total epidemic size. Finally, they evaluate the risks in the worst-case scenario consisting in a very long decay phase.

Soubeyrand aims at combining pathogen genomic data with spatial and temporal data to estimate the transmission tree of an infectious disease in the host population. He presents a model that combines an individual-based semi-Markov SEIR model for the spatio-temporal dynamics of the pathogen, with a Markovian evolutionary model for the temporal evolution of genetic sequences of the pathogen. He proposes a new algorithm for Bayesian inference on the parameters and assesses its performances through simulations.

Finally, Madouasse *et al.* illustrate the use of statistical methods for various types of data in veterinary epidemiology by considering the emergence of bluetongue virus in cattle in 2006 as a case study. They show how to estimate the impacts of the disease on milk production and on reproduction as well as the impact of the vaccination on the reproduction. Moreover they present syndromic surveillance methods for the detection of vector-borne emerging diseases and assess their performances on the basis of simulated data. The authors also discuss challenges and problems facing veterinary epidemiology that would benefit from deepening the collaborations between veterinary epidemiologists and statisticians.

We are grateful to all authors for their contributions, to referees for their careful reading and to the editor Gilles Celeux for his invitation to publish this special issue.