

Public health related decisions concerning infectious diseases are characterized by the use of complex mathematical models as an essential tool in understanding the dynamics of infectious diseases and designing intervention strategies. Efficient modelling and inference procedures for learning the model parameters from data are of central interest. The aim of the conducted research is to synthesize data and techniques from a number of disciplines in order to solve this problem. New and efficient statistical machine learning methods, namely Hamiltonian Monte Carlo (HMC) and Variational Inference (VI) are implemented for several deterministic and stochastic epidemic models using the relatively new freely available software package Stan. Stan is the first software which implements HMC and VI routines while offering built-in solvers for systems of ordinary differential equations (ODEs). This makes Stan a particularly attractive candidate tool for fitting models based on ordinary differential equations appearing both in epidemiology and economics. In the paper "[Contemporary statistical inference for infectious disease models using Stan](#)" (with E. van Leeuwen, O. Ratmann, M. Baguelin and N. Demiris) we explore how Stan could be used to fit mathematical models to infectious disease count data and offer a brief description of the most important features of Stan's inference algorithms so the reader can get familiar with the tools. The latest official Stan documentation can be found <https://mc-stan.org/> and <https://epidemiology-stan.github.io/>.

More than a year into the current SARS-CoV-2 pandemic, action plans have been implemented around the globe in order to reduce transmission. The reproduction number has been found to respond to public health interventions changing throughout the pandemic waves. However, the actual global burden of SARS-CoV-2 remains unknown due to severe under-ascertainment of cases. The use of reported deaths has been pointed out as a more reliable source of information, likely less prone to under-reporting. Given that daily deaths occur from past infections weighted by their probability of death, one may infer the true number of infections accounting for their age distribution, using the data on reported deaths. In a joint work with Kostas Kalogeropoulos (London School of Economics and Political Science) and Nikolaos Demiris (Athens University of Economics and Business), we adopt this framework and assume that the transmission dynamics generating the total number of underlying infections can be described by a continuous time transmission model expressed through a system of non-linear ordinary differential equations, where the transmission rate and consequently the reproduction number are stochastic. We model the transmission rate as a diffusion process, with different volatility values for each pandemic wave, allowing to reveal both the effect of control strategies and the changes in individuals behavior. We study the case of 6 European countries and estimate the time-varying reproduction number ( $R_t$ ) as well as the true cumulative number of infected individuals using the No-U-Turn sampler variant of Hamiltonian Monte Carlo using the Stan software. As we estimate the true number of infections through deaths, we offer a more accurate estimate of  $R_t$ . We also provide an estimate of the daily reporting ratio and discuss the effects of changes in mobility and testing to the inferred quantities.

Anastasia Chatzilena states that: "Part of this research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY)".