

# DSS Statistics Seminar

March 12, 2021, 12:00

<https://uniroma1.zoom.us/j/86881977368?pwd=SWRFcVFjMDZTa0lXZk05TE1zNm5adz09>

Passcode: 432940

Modelling dirty surveillance data  
during the Italian epidemic  
outbreak of SARS-CoV-2

## ***StatGroup-19***

***P. Alaimo di Loro, F. Divino, A. Farcomeni, G. Jona Lasinio, G. Lovison, A. Maruotti, M. Mingione***

We briefly discuss the main epidemiological features of SARS-CoV-2 one year into the pandemic, giving a short account of Italy's public data and the main limits of lay analyses.

We then discuss an accurate method for short-term forecasting ICU occupancy at the local level. Our approach is based on an optimal ensemble of two simple methods. We combine a generalized linear mixed regression model which pools information over different areas and an area-specific non-stationary integer autoregressive methodology. Optimal weights are estimated using a leave-last-out rationale.

Daily predictions between February 24th and November 27th, 2020 have a median error of 3 beds (third quartile: 8) at a regional level, with coverage of 99% prediction intervals that exceed the nominal one.

We present a different method based on a modified non-linear GLM for each incidence indicator, including the potential effect of exogenous variables, based on appropriate distributional assumptions and a logistic-type growth curve. That allows us to accurately predict essential characteristics of the epidemic (e.g., peak time and height).

Eventually, some considerations on the evolution of the non-linear GLM approach will be discussed.



**SAPIENZA**  
UNIVERSITÀ DI ROMA