G. Tunnicliffe–Wilson's contribution to the Discussion of 'New tools for network time series with an application to COVID-19 hospitalisations' by Nason et al.

Granville Tunnicliffe Wilson, Department of Mathematical Sciences, Lancaster University

It is a real statistical challenge to elicit, from historic records, the dependencies between a large collection of time series. To meet this challenge the GNAR model considered in this paper takes advantage of the spatial structure of the series. It does this by using as predictors of each series, specific linear combinations of neighbouring series in a network. The authors provide tools for identifying the spatial extent of the networks and the useful lags for prediction. Their application to COVID-19 hospitalisations is most valuable for illustrating the model and its effectiveness.

Although the methodology is well illustrated by figures and plots the presentation would be better informed by graphs of some the series, so I show plots of two of the bed occupancy series below. I ranked the series according to their mean value; the first series has the highest mean, the second has rank 35 out of 140, starting from the lowest mean. The value of these is firstly to show that the integer nature of the observations necessarily adds to the prediction error, and substantially so for the series with low means. This is accentuated by use of the logarithmic transformation (after adding one) which generally leads to large variability at low levels, except when the series is constant at zero for several weeks. This helps understand why the naïve predictor, using just the previous value, is quite competitive. It also helps understand the value of the GNAR model, because the combination of related network series extracts a common signal, reduces the noise, and produces a more valuable predictor. The reduction in the MSPE of the best GNAR model over the naïve predictor may seem modest, but it indicates a real improvement in prediction of the underlying movement of each series.

But I must comment on the stated MSPE of the AR(1) in this model; my own computations yield an MSPE of 7.4. I also find that a similar assessment at an earlier point where the series is locally trending, shows a double exponential smoother model, designed to track changing level and trend, outperforming the AR(1) and Naïve predictors.

My final comment is that the five series with highest means, from Manchester, Birmingham, and three London hospitals, show strong low frequency coherency. We can expect any univariate predictor to be outperformed by a well-constructed and parsimonious model, such as the GNAR, which uses information from related series.

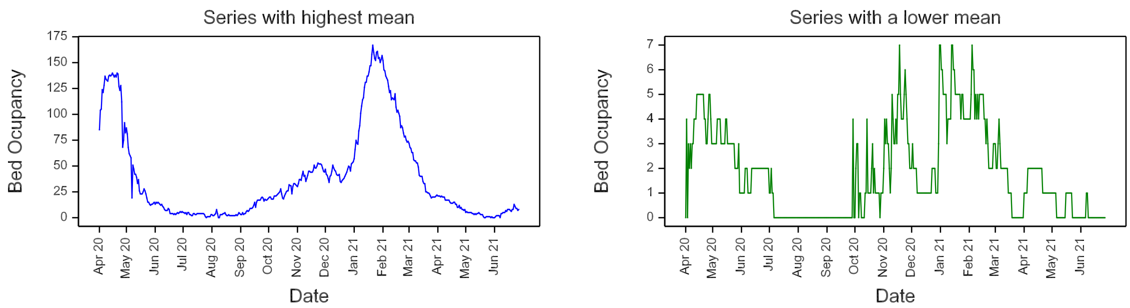

Figure 1: Plots of time series of Mechanical Ventilation Bed Occupancy.