

# Who's in the Waiting Room? Modelling Multivariate Time Series of Counts of Patients to Hospital Emergency Departments

Sarah Bolt<sup>1,\*</sup>, Ross Sparks<sup>1</sup>, James Lind<sup>2</sup>

1. CSIRO Mathematics, Informatics and Statistics, North Ryde, NSW, Australia

2. Gold Coast Hospital, Southport, QLD, Australia

\*Contact author: [sarah.bolt@csiro.au](mailto:sarah.bolt@csiro.au)

**Keywords:** Multivariate, Times Series, Health Informatics.

While predictive tools are already being implemented to assist in forecasting the total volume of patients to Emergency Departments [Jessup et al. (2010)], early detection of any changes in the types of patients presenting would help authorities to manage limited health resources and communicate effectively about risk, both in a timely fashion. But before we are able to detect changes we must understand and model the expected counts of presentations for all possible subgroups. For example, we need to forecast the expected number of patients on a given day at a particular hospital, with a particular disease, of a particular age, sex, etc.

So our objective was to model this large collection of interdependent time series. This problem presented a number of issues including the sheer number of them and the fact that many of these time series had very low counts and displayed overdispersion relative to a Poisson model. Furthermore, since the goal of these models was to model away known behaviours, we had to incorporate effects including significant seasonality, day of the week effects and some strongly interacting variables.

The method we present here divided the problem into two different components. The first used a regression approach to model time series of aggregated counts. The second allocated proportions of these counts to subgroups using a binary regression tree analysis. This method thus drew together *R* functionality from two different areas:

- Analysis of times series of counts using `glm {stats}` [R Development Core Team (2010)], `glm.nb {MASS}` [Venables and Ripley (2002)] and `gamlss {gamlss}` [Rigby and Stasinopoulos (2005)]. The latter function allowed us to model counts with a negative binomial distribution and a modelled dispersion parameter.
- Allocation of counts to subgroups using the poisson regression tree approach of `rpart {rpart}` [Therneau et al. (2010)].

## References

- Jessup, M., M. Wallis, J. Boyle, J. Crilly, J. Lind, D. Green, P. Miller, and G. Fitzgerald (2010). Implementing an emergency department patient admission predictive tool. *Journal of Health Organization and Management* 24, 306–318.
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Rigby, R. A. and D. M. Stasinopoulos (2005). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics* 54, 507–554.
- Therneau, T. M., B. Atkinson, and B. D. Ripley. (2010). *rpart: Recursive Partitioning*. R package version 3.1-48.
- Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S* (Fourth ed.). New York: Springer. ISBN 0-387-95457-0.