Forecasting COVID-19 Counts at a Single Hospital: A Hierarchical Bayesian Approach

Alexandra Hope Lee*, Panagiotis Lymperopoulos*, Joshua T. Cohen[1], John B. Wong[2], and Michael C. Hughes*

Tufts UNIVERSITY

* *Dept. of Computer Science, Tufts University*
1 Center for Evaluation of Value and Risk, Tufts Medical Center
2 Division of Clinical Decision Making, Tufts Medical Center

Python code and paper on GitHub:
github.com/tufts-ml/single-hospital-count-forecasting

Goals

What: Predict next 1-3 weeks of COVID-19 occupancy counts at a specific hospital, given univariate time series of past counts

Why: Help local stakeholders predict and respond to future demand

How:
- Probabilistic models and Bayesian analysis, because proper estimation and communication of uncertainty is important.
- Hierarchical modeling of multiple hospitals in the same region to improve forecasts.

Model

Latent real values f
Observed counts y

$$p(f_{1:T}, y_{1:T}) = p_\alpha(f_{1:T}) \cdot \prod_{t=1}^{T} p_\lambda(y_t \mid f_t)$$

- Latent autoregressive process (GAR)

$$\prod_{t=1}^{T} \mathcal{N}(f_t | \beta_0 + \sum_{w=1}^{W} \beta_w f_{t-w}, \sigma^2)$$

- Latent Gaussian process (GGP)

$$\mathrm{GP}(m_\alpha(t), k_\alpha(t, t'))$$

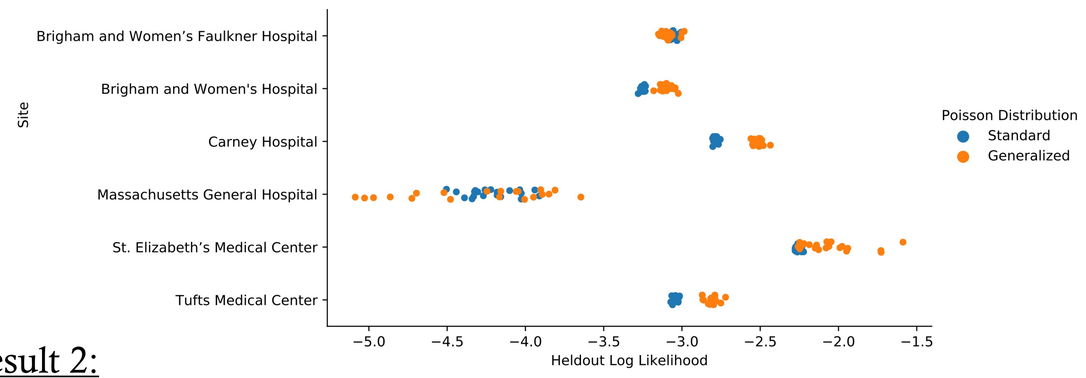
- Standard Poisson
- Generalized Poisson *(Consul & Famoye, 1992)*
  More flexible with learnable dispersion
  New code for PyMC3

Multi-site: share likelihood parameters
HMC MCMC in PyMC3 *(Salvatier et al. 2016)*

Retrospective Evaluation

Look back at early summer 2020 at 8 sites in MA, 10 sites in UK

Result 1: Generalized Poisson better predictions than standard Poisson



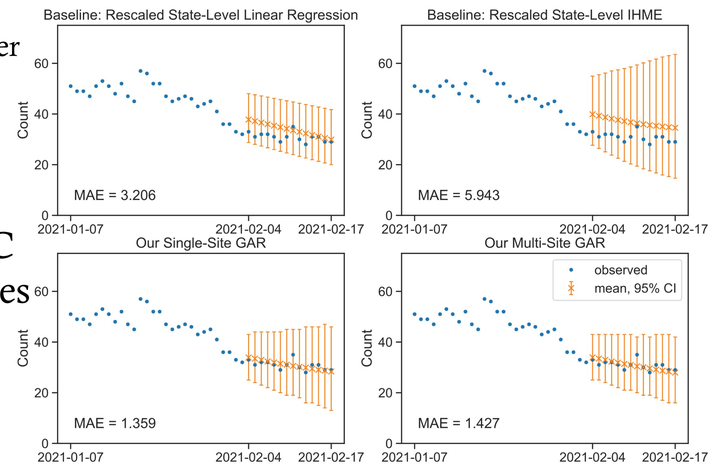
Result 2:
GAR autoregressive model more reliable than Gaussian process (GGP)
Multi-site model best for 6/8 nearby sites in Suffolk County, MA
Multi-site less dominant but still competitive on 10 sites across England

Hospital Site (MA, USA)	**Single-Site GGP**	**Single-Site GAR**	**Multi-Site GAR**
Beth Israel Deaconness	-3.153 ± 0.005	-3.121 ± 0.017	-3.010 ± 0.011 *
Boston Medical Center	-3.376 ± 0.002	-3.454 ± 0.025	-3.296 ± 0.014 *
Brigham & Women's Faulkner	-2.719 ± 0.002	-2.684 ± 0.010	-2.601 ± 0.008 *
Brigham & Women's Hospital	-3.118 ± 0.001	-2.978 ± 0.009*	-3.044 ± 0.016
Carney Hospital	-3.072 ± 0.008	-2.721 ± 0.105	-2.564 ± 0.090 *
Massachusetts General	-3.777 ± 0.020	-3.468 ± 0.018 *	-3.450 ± 0.017 *
St. Elizabeth's	-1.980 ± 0.007	-1.312 ± 0.120 *	-1.598 ± 0.007
Tufts Medical Center	-2.937 ± 0.005	-2.822 ± 0.010	-2.728 ± 0.007 *
Hospital Site (UK)	**Single-Site GGP**	**Single-Site GAR**	**Multi-Site GAR**
Barts Health	-4.730 ± 0.058 *	-6.727 ± 0.221	-4.842 ± 0.078
Chelsea & Westminster	-2.937 ± 0.009 *	-3.141 ± 0.030	-5.929 ± 0.086
Imperial College Healthcare	-4.017 ± 0.019	-3.063 ± 0.178 *	-3.976 ± 0.110
King's College Hospital	-3.012 ± 0.003	-2.749 ± 0.020	-2.695 ± 0.005 *
London North West University	-2.829 ± 0.008	-2.488 ± 0.014 *	-2.572 ± 0.009
Manchester University	-12.335 ± 0.647	-12.735 ± 0.937	-6.152 ± 0.167 *
North Middlesex University	-1.593 ± 0.008 *	-1.677 ± 0.022	-1.703 ± 0.008
Nottingham University	-3.443 ± 0.018	-3.324 ± 0.066	-2.950 ± 0.032 *
University Hospitals Birmingham	-4.385 ± 0.005 *	-4.571 ± 0.019	-4.583 ± 0.031
Univ. Hospitals of North Midlands	-3.188 ± 0.002	-2.953 ± 0.033	-2.829 ± 0.016 *

Prospective Evaluation

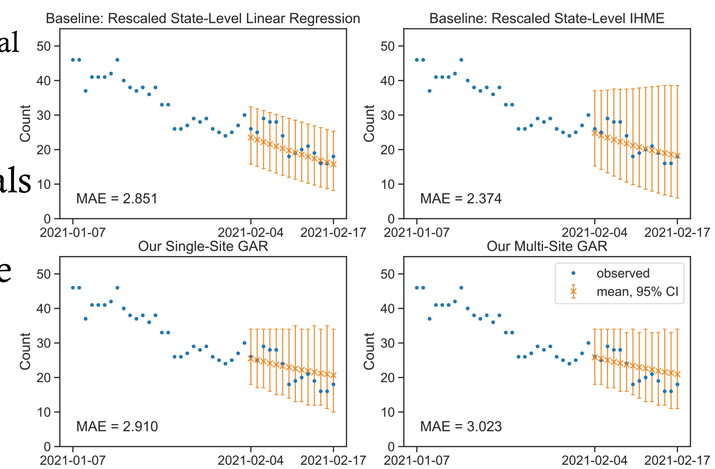
In early Feb. 2021, we made true 2-week-ahead forecasts for all 3 hospitals managed by Wellforce, a health system in MA. We compared our methods to two baselines currently in use to help local stakeholders react to demand.

Tufts Medical Center



Result 3: Absolute error reduced
- MAE reduction of 1.75 at TMC
- MAE comparable at 2 other sites

Melrose-Wakefield Hospital



Result 4: Better uncertainty intervals
- IHME baseline too wide
- Linear baseline does not become more uncertain further in future