CoDatMo:
Creating a Common Epidemiological Framework with Bayesian Models

Breck Baldwin, PhD
fbb2116@columbia.edu

Who

- Breck Baldwin: Project Manager Laplace Lab @Columbia statistics
- Andrew Gelman: @Columbia statistics
- Simon Maskell: PI Big Hypotheses Lab at U of Liverpool
- Conor Rosato, Lee Delvin, John Heap, Philip Clemson, Allesandro Varsi, Matthew Carter, Robert Moore, Vincent Beraud, Al Philips, Kelli Casidy, Jose Storypoli, Alessandra Pellini, Elias Noda, Bob Carpenter

How

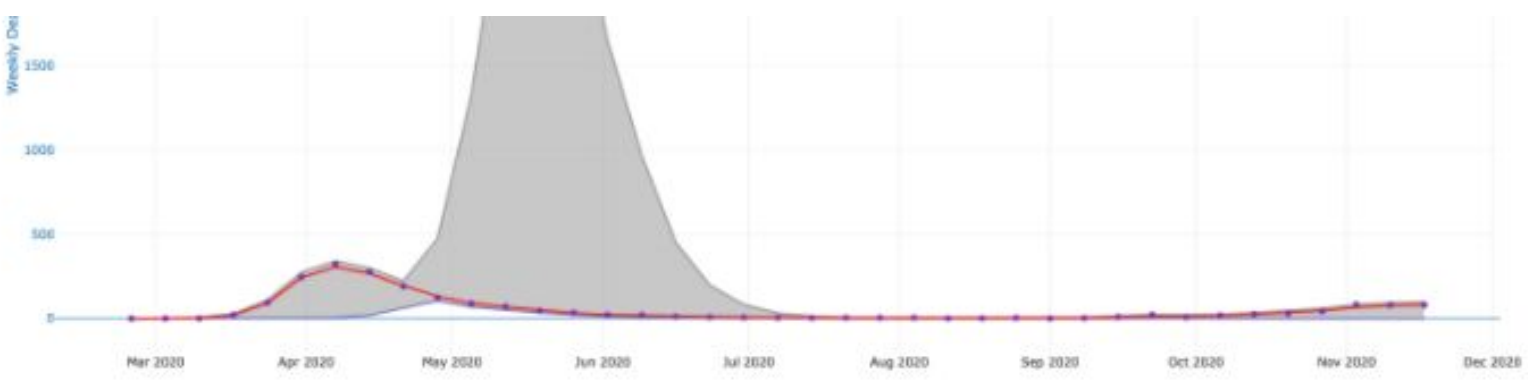
- UKRI Big Hypotheses EP/R018537/1
- NSF RAPID grant 2055251

Outline

- Overview CoDatMo
- Brief intro to Bayesian Modeling and Stan
- Go over some repositories
- Resources

CoDatMo Goals and History

1. August 2020:
    a. Stan--a popular language for modeling COVID-19 epidemiology ~ 20 models
    b. The UK Joint Biosecurity Center (JBC) uses many models but similarities/differences
    c. Liverpool/Columbia decide to create an online resource for comparative COVID modeling
        i. Improve the models
        ii. Share ideas
2. November 15, 2020: NSF RAPID grant 2055251 for COVID awarded
3. December 2, 2020: https://github.com/codatmo up and running
    a. Model example
    b. Data
4. Submissions:
    a. December 2, 2020: Liverpool
    b. December 8, 2020: Ye Old Lancaster
    c. December 20, 2020: Simple SIR
    d. …..Rethinking appeal of CoDatMo to external contributors
    e. March 14, 2020: UNINOVE Sao Paulo
5. April 14, 2020: This presentation



Death incidence in nhs_sandwell_and_west_birmingham_ccg

Welcome to the CoDatMo site

CoDatMo (COvid DATa MOdel) exists to replicate and host COVID-19 models written in the Bayesian modeling languages like Stan or PyMC. The goals include:

- Reproduce COVID-19 models with an eye to:
  - Get the word out on important COVID-19 models by offering well documented examples approachable by statiticians, programmers and researchers.
  - Increase the impact of the original researchers and get them additional exposure.
  - Help researchers improve their models via the replication process
- Provide a common language for implementing various models.
- Facillitate information exchange between model developers.
- Provide public access to models.

CombineR

R BSD-3-Clause 0 0 0 0 Updated 4 hours ago

UNINOVE_Sao_Paulo

Covid Modeling work at UNINOVE Brazil

HTML BSD-3-Clause 1 2 0 0 Updated 3 days ago

Data

Collection of open-source data sources for the United Kingdom and the rest of the world. To be used for testing of the various models.

HTML BSD-3-Clause 0 0 0 0 Updated 17 days ago

Liverpool

HTML BSD-3-Clause 1 0 0 0 Updated 17 days ago

Top languages

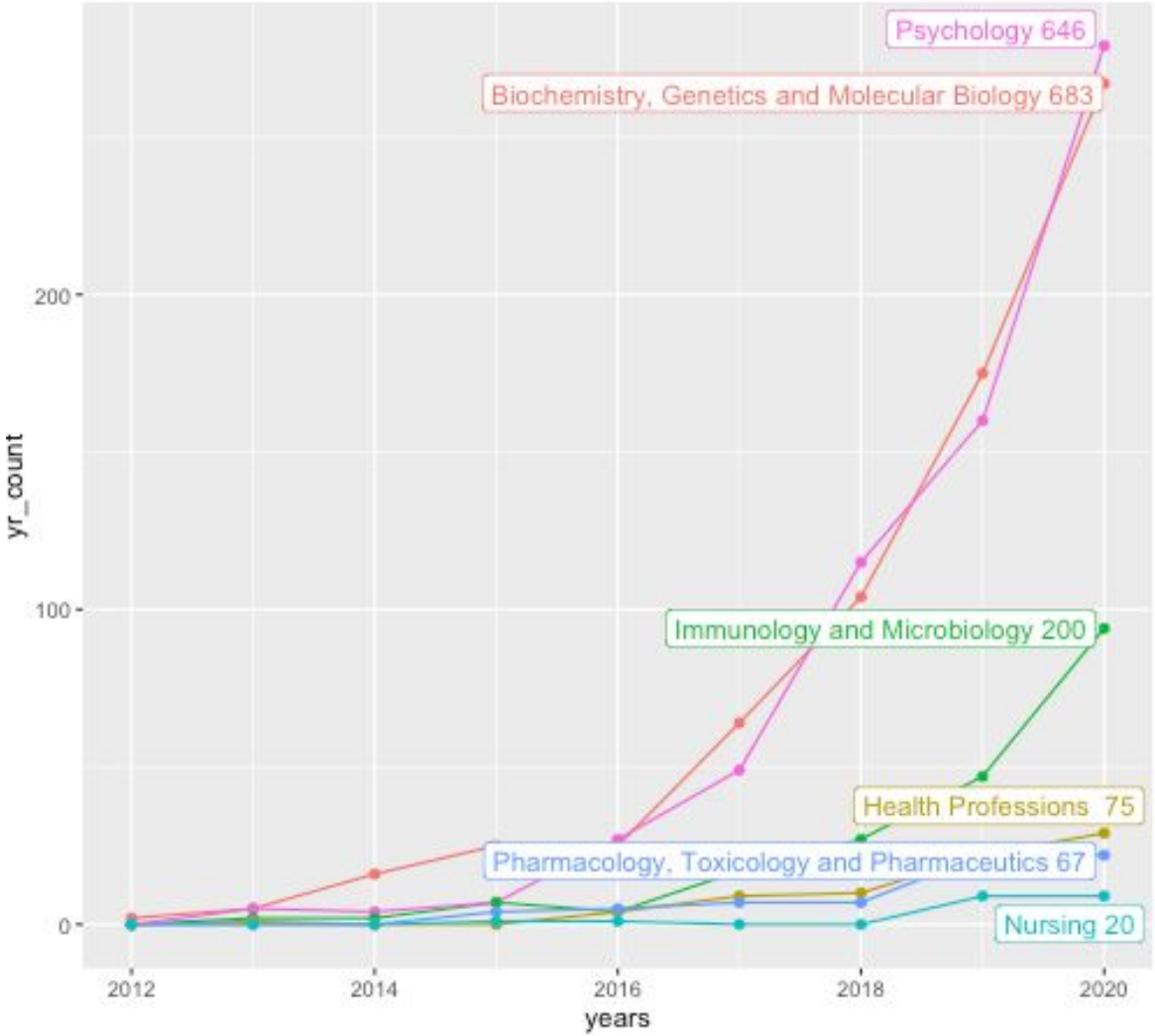
HTML R Stan

People

Invite someone

Over 50 models for COVID-19 written in Stan: Why?

- Hierarchical models
- Noise tolerant models with small to medium data sets
- Lots of room for integration of prior knowledge
- Can be very interpretable models to support counterfactual reasoning etc…
- Very robust uncertainty handling
- Huge effort on inference safety
  - Bayesian workflow
  - Inference diagnostics
  - Self monitoring gradient crawler
- Generative models
  - Mechanistic
  - Human interpretable
  - Simulation


Psychology 646
Biochemistry, Genetics and Molecular Biology 683
Immunology and Microbiology 200
Health Professions 75
Pharmacology, Toxicology and Pharmaceutics 67
Nursing 20
yr_count
200
100
0
2012
2014
2016
2018
2020
years

Deep Learning

Repos in CoDatMo

https://codatmo.github.io/Simple_SIR/

https://github.com/codatmo/UNINOVE_Sao_Paulo

Resources

- Codatmo project: https://codatmo.github.com
  - Help us...
  - Let us help you...
- Slides: https://github.com/codatmo/documentation/blob/main/talks/CovidSymp_4_16_2021.pdf
- Modern Bayesian modeling packages:
  - Stan: https://mc-stan.org
    - R, Python, Scala, ... https://mc-stan.org/users/interfaces/
    - High level interfaces in lme4 tradition (R only)
      - brms: https://paul-buerkner.github.io/brms/
      - RStanArm: https://mc-stan.org/rstanarm/
  - PyMC Python HMC/NUTS: https://pypi.org/project/pymc/
  - TensorFlow/PyTorch
- Statistical Rethinking, Second Edition, by Richard McElreath
  - Youtube lectures https://youtube.com/playlist?list=PLDcUM9US4XdNM4Edgs7weiyIguLSToZRI
  - brms implementation of 1st Edition at: https://bookdown.org/content/4857/
- Bayesian Statistics for the Social Sciences GR5065 in http://www.qmss.columbia.edu/:
- Contact me: fbb2116@columbia.edu