Probabilisitic epidemic forecasting using probability generating functions, and its robustness to data quality, error, biases, and noise

The University of Vermont

Mariah C. Boudreau, Christopher M. Danforth, & Laurent Hébert-Dufresne

Introduction

- Probability generating functions (PGFs) are analytical and probabilistic tools that describe random networks [1]
- PGFs predict the final outbreak size of stochastic disease transmission trees
- This research explores the effect of error on PGFs

Methods

To assess the sensitivity of the PGF, multiplicative perturbations were added to the model.

$$G_1(x) = \sum_{k=0}^{\infty} p_k(1 + ERROR)x^k$$

- Error is from a normal distribution ($\mu = 0, \sigma^2 =$ varied)
- Coefficients are the probability of k potential transmission(s)
- Simulations were run to access the final outbreak size over various network conditions

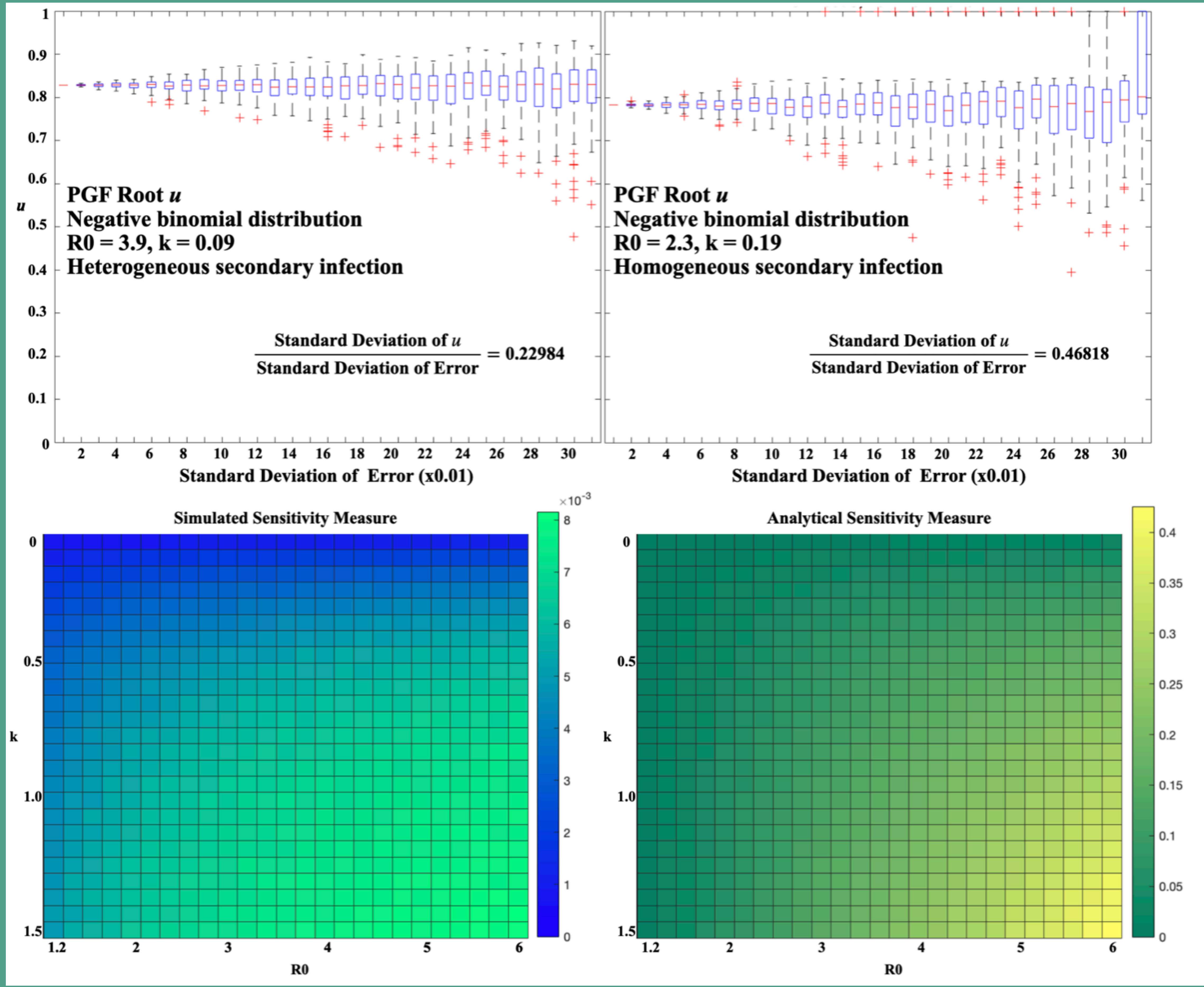
- Sensitivity measures were calculated for the root u for the self-consistent equation

$$u = G_1(u)$$

- Where the final outbreak size is equal to S [3]

$$S = 1 - u$$

Results



Top: The distributions of the PGF roots for two network conditions (R0 = 2.3, k = 0.19 and R0 = 3.9, k = 0.09) exhibits the affect error has on the system. *Bottom left:* The heat map exhibits the change in root u over the change in the coefficients, giving a sensitivity measure for each R0 and k value. *Bottom right:* The heap map exhibits analytical sensitivity measure of the root u described by Winkler giving a sensitivity measure for each R0 and k value.

Analysis

- Sensitivity measures for polynomial roots are detailed in Winkler's paper [2]
- Analysis through probability theory
- Componentwise Sensitivity is

$$\gamma_0 = \frac{E(\Delta u)}{E(\Delta p_{k_c})} = \left| \frac{1}{G_1'(u)} \right| \frac{||\sum p_k u^k||_2}{u||p_k^{-1}||_1}$$

- Ratio of the expected error in the root over the expected error in the coefficients

Discussion

- Homogeneous network appears more sensitive to error, given simulated analysis
- Heterogeneous networks with lower R0 value appear more robust to added error
- Simulated results agree with the analytical results to varying degrees

References

1) Newman, MEJ, Strogatz, SH, and Watts, DJ. "Random graphs with arbitrary degree distributions and their applications." *Physical Review E* 64.2 (2001): 026118.
2) Winkler, J.R. "A statistical analysis of the numerical condition of multiple roots of polynomials." *Computers & Mathematics with Applications* 45. 1-3 (2003): 9-24.
3) Lloyd-Smith, J, Schreiber, S, Kopp, P, *et al*. "Superspreading and the effect of individual variation on disease emergence. *Nature* 438, 355-359 (2005).