When Local Virus Outbreaks Become a Global Health Concern – How to Detect Them Earlier Than Witnessed for Ebola in 2014

Irene Hueter, Columbia University, New York, Novartis IIS NIBR / GCE Solutions, IL (BASS sponsor)

November 3, 2015

Irene Hueter, Ph.D.

Email: irene.hueter@columbia.edu
irene.hueter@gcesolutions.com

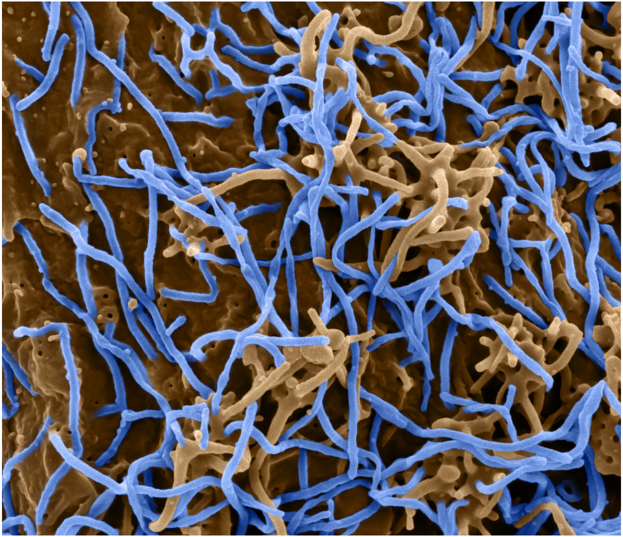
Disclaimer

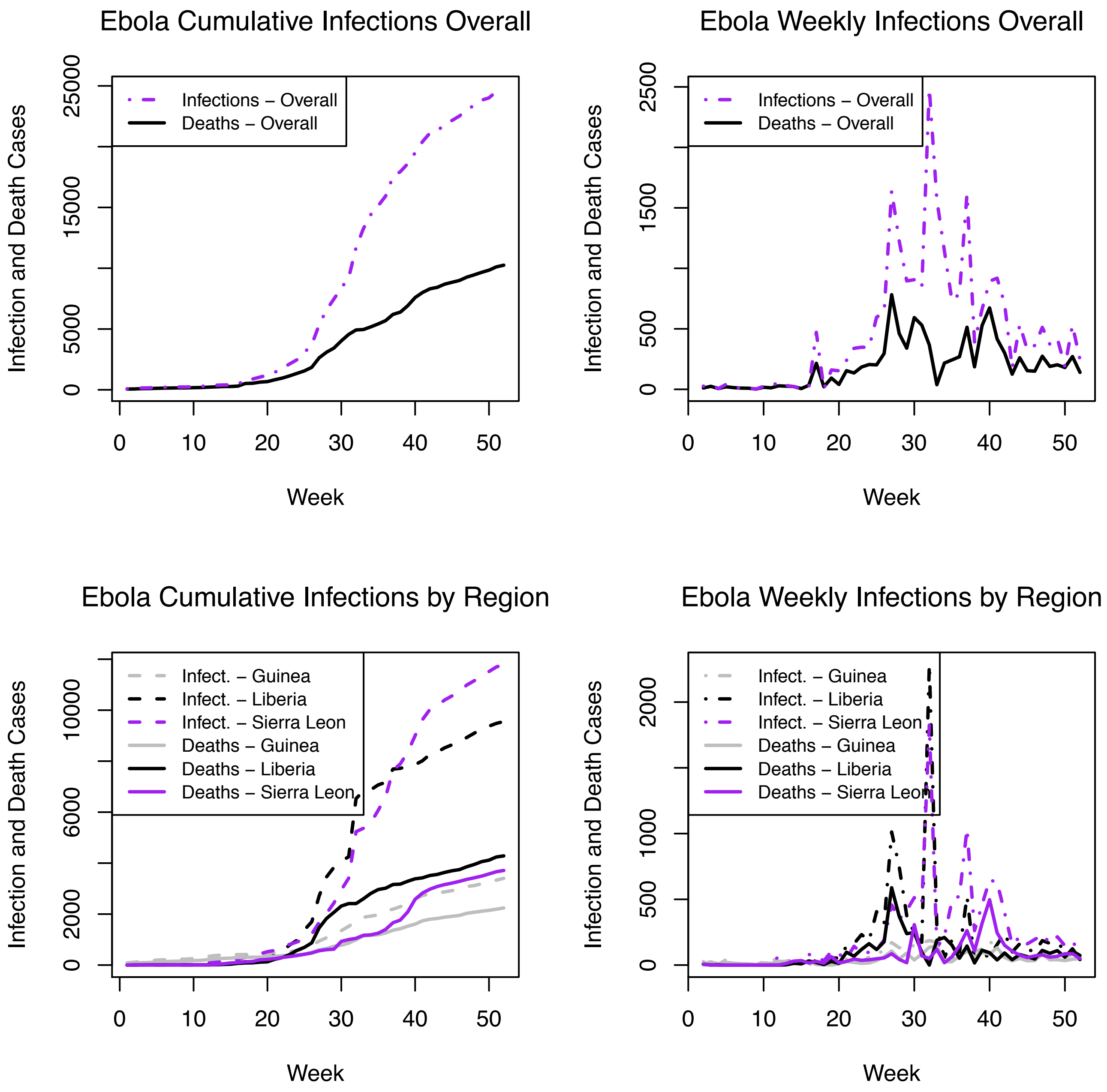
- Presentation is based on research and two preprints by speaker submitted for journal publication.
- Contents of presentation represents confidential research information and should not be shared or disseminated except by the author or upon the author's approval.
- This presentation document or parts of it including figures must not be disseminated or posted in a publicly accessible site.

Overview

One Year Ago...

Ebola epidemic was at its peak and media coverage was intense.




Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
5000
15000
25000
0
10
20
30
40
50
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
500
1500
2500
0
10
20
30
40
50
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
2000
6000
10000
0
10
20
30
40
50
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1000
2000
0
10
20
30
40
50
Week

Cost to Halt Outbreak Out of Control

Quotes from News Postings:

- *'U.S. may spend up to $1 billion fighting deadly disease, Ebola, Obama administration says on September 16, 2014'*
- *'The administration has already spent $175 million responding to outbreak.'*
- *'Planned use of resources of U.S. military to establish up to 17 treatment centers in Liberia and train as many as 500 healthcare workers a week in region to cope with outbreak.'*

Cost to Halt Outbreak Out of Control

- *'International Medical Corps, a non-profit group working in West Africa, estimates on November 2, 2014 that it will cost $1.6 billion over next 6 months to bring disease under control.'*
- *'Ebola's economic devastation worsened in West Africa, World Bank said in December 2014, predicting gross domestic product would shrink this year in Sierra Leone and Guinea.'*
- *'Connecticut hospitals have spent more than $5 million on preparations to deal with possible Ebola patients, lawmakers were told Monday (Nov 1, 2014).'*

Economic Cost of 2014 Ebola Outbreak

Quotes from News Postings:

- *'If Ebola continues to spread further in Africa, it could cost as much as \$32.6 billion by end of 2015, ...estimated in October.'*
- *'We simply must find resources required, no matter cost, to get to zero cases as soon as possible.'*
- *'Defeating Ebola now will cost billions – but it will spare rest of world from spread of virus, save lives in countries, save money over long term, and help countries rebuild their economies (November 2014).'*
- *'Government still number-crunching with IMF, but reckons Ebola will shave over 2% of 2014 growth rates.'*

Economic Cost of 2014 Ebola Outbreak

- *'Mostly because of damage done to mining, agriculture and service industries, as investors evacuate foreign workers, borders close, and international flights are suspended.'*
- *'Bread-basket regions are under quarantine, making agricultural trade impossible.'*
- *'On December 17, UN and World Food Program estimated that 120,000 Sierra Leoneans have become food insecure as a result of Ebola, meaning they neither have food they need nor are able to buy it.'*

Drastic Interventions on Outbreak Occurred

According to Various News Reports and Press Releases:

Drastic interventions

- in large parts were initiated in late September of 2014,
- reportedly included an international response that comprised over 62 countries,
- involved contributions $>$ \$2 billion, a few thousand military troops who arrived in Liberia to establish 15 Ebola treatment centers and to train around 1500 healthcare staff.

Drastic Interventions on Ebola Outbreak

A Few Observations:

- Drastic interventions to halt outbreak started to occur in late September and October of 2014.
- Exponential growth of 2014 Ebola infection or death cases became irrefutable in early July of 2014.
- It was obvious in March and April of 2014.
- Ebola virus continued to spread at a rate in the supercritical regime until November of 2014.
- Interventions began to show an effect in November of 2014.

Consequences of Delayed Intervention

A Few Observations (Cont'd):

- Intervention and economic costs are directly related to observed weekly number of infection cases and also depend on duration of outbreak.
- The higher the weekly number of infection cases, the longer the waiting time to reach zero or a few infection cases.
- Delays of needed interventions by several weeks allow peak number of weekly infections to further increase at supercritical rate.

Is Intervention Effect Sufficient?

- In February of 2015, weekly number of infection cases have returned to levels last recorded in August of 2014.
- This time infection rate is decreasing not increasing.

Question: Had intervention sufficient effect to stabilize weekly number of infection cases and eventually bring them to zero?

Two Problems of Interest to Address

1. Early detection and accurate prediction of magnitude of outbreak several months before it spins out of control

2. Timely assessment as to whether an intervention has sufficient impact to stabilize and eventually end it

Four-decades long history of Ebola and Marburg in humans teaches us that filovirus outbreaks recur.

A Look Backwards

Number of Infection and Death Cases to Date?

Ebola Data Set Used

- Weekly numbers of infections over time overall & by region
- Time period: March 24, 2014 through February 25, 2015
- Data Source: website www.cdc.gov, Centers for Disease Control and Prevention
- Infection cases: 'suspected,' 'probable,' and 'confirmed'
- Approximately weekly infection cases: data recorded and reported have not been completely regular

Let's Look at a Few Numbers of Outbreak Now

In Guinea, Sierra Leon and Liberia, **on September 27, 2015:**

- $\sim$ 28,319 infection cases (suspected, probable, and confirmed)
- $\sim$ 11,296 death cases
- $\sim$ 4–70 new infection cases per week (Guinea, Sierra Leon)
- $\sim$ 1-3 new death cases per week (Guinea, Sierra Leon).

Week of October 8th '15 is first week since outbreak was declared in March '14, when no new case was reported by WHO.

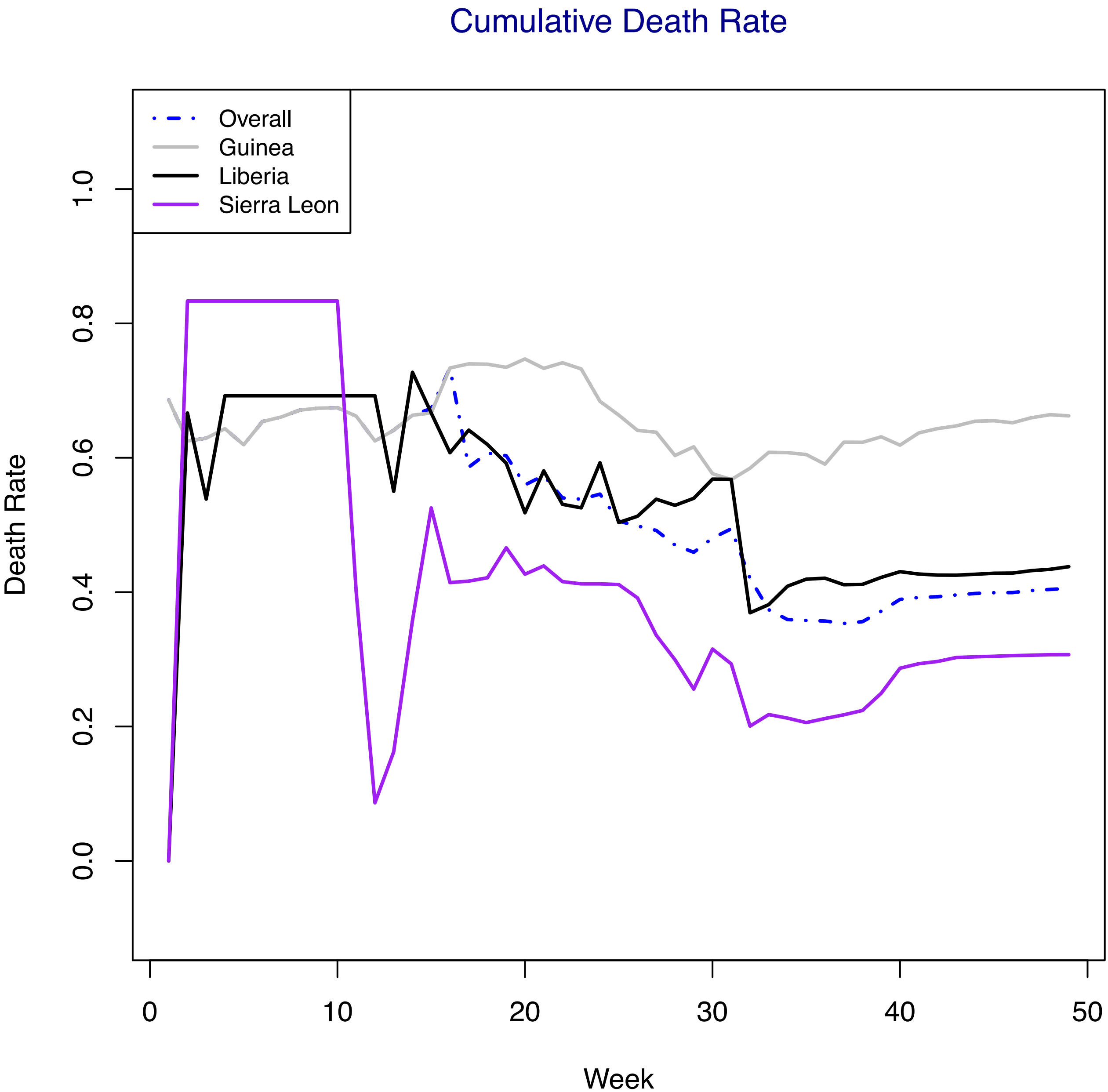
Let's Look at a Few Numbers of Outbreak Now

In Guinea, Sierra Leon and Liberia, **on October 17, 2015:**

- $\sim$ 28,513 infection cases (suspected, probable, and confirmed)
- $\sim$ 11,313 death cases
- Week of October 8th '15 is first week since outbreak in March '14, when no new case was reported by WHO.
- CDC reported: 48 new infection cases, 1 new death case, 3 new laboratory cases since October 6 '15.

Some Facts on Outbreak

- 2014 outbreak in Guinea, Liberia, and Sierra Leon was largest Ebola outbreak ever.
- Ebola is one of world's most deadly diseases that can kill majority of those infected within days.
- Initial death rate $\sim$ 70% (during March through July of 2014) steadily decreased to $\sim$ 36% (December 2014).
  Death rate = (total nr. of deaths)/(total nr. of infections)
- Outbreak's first suspected case, two-year old child who died in December 2013, after being sick for four days, in Guéckédou, a Meliandou village.
  So did child's sister, mother, grandmother, and village midwife after hospitalization shortly after.

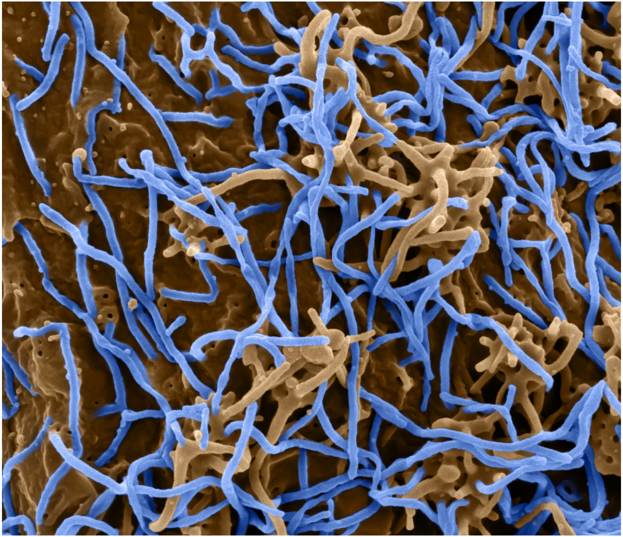


Cumulative Death Rate
Overall
Guinea
Liberia
Sierra Leon
Death Rate
0.0
0.2
0.4
0.6
0.8
1.0
Week
0
10
20
30
40
50

Brief History on Ebola and Related Filoviruses

- Ebola belongs to family of viruses, known as filoviruses (means 'thread virus' in Latin).
- Filoviruses look like strands of tangled rope or hair.
- Family comprises virus types called Ebola Zaire, Ebola Sudan, Marburg, and Ebola Reston among others.
- Some form of Ebola can travel through air via aerosol of droplets. Was observed in monkeys.
- Current Ebola strain that emerged in West Africa is a different strain from those mentioned above.

Source: Book, NYT Bestseller *'The Hot Zone: A Terrifying True Story' by Richard Preston,1994*



Brief History on Ebola and Related Filoviruses

- **Marburg** was discovered in Marburg, Germany, in 1967, in a factory which produced vaccines using kidney cells from African green monkeys, imported from Uganda.
- Shortly afterwards, Marburg agent jumped species and emerged in human population of city Marburg and caused first human death in 1967.
- Kills 1 out of 4 humans infected, thus, death rate $\sim$ 25%.
- In 1980, Marburg virus re-emerged in human population in Kenya near Mount Elgon, a volcano rising to 14,000 feet. Was traced back to persons who entered Kitum Cave.

Brief History on Ebola and Related Filoviruses

- First case of **Ebola Sudan** happened in 1976.
- Killed 1 out of 2 humans infected, thus, death rate $\sim$ 50%.

Brief History on Ebola and Related Filoviruses

- First known emergence of **Ebola Zaire** occurred in 1976, two months after the start of the Sudan emergence.
- Simultaneously emerged in 55 villages near Ebola River in northern Zaire, now Democratic Republic of Congo. Appeared to come out of nowhere.
- First human case of Ebola Zaire has not been identified. It is conjectured that a blood-to-blood contact in rain forest enabled virus to move into the human world.
- Killed 9 out of 10 humans infected, thus, death rate $\sim$ 90%.

Brief History on Ebola and Related Filoviruses

- **Ebola Reston**, which looks similar to Ebola Zaire, emerged in a 'monkey house' in Reston, northern Virginia, in 1989, which was a quarantine station for imported lab animals.
- 100 monkeys shipped from Philippines died.
- Was spreading through air in monkeys but not for humans.
- Humans infected with this Ebola strain did not die. Strain distinguished between monkeys and humans.

Outbreak Data Set

- During December 2013-February 2014, there were only scattered news reports about Ebola cases and onset of outbreak.
- By February and March 2014, it became obvious that outbreak was evolving, and differently from previous smaller and local ones.
- Data were not systematically collected and reported until March 2014.
- Data Source: website www.cdc.gov, Centers for Disease Control and Prevention
- Soon it became evident that infection and death cases were increasing exponentially with time and disease had spinned out of control.

Some Facts on Outbreak

In March 24, 2014:

- $\sim$ 86 infection cases (Guinea)
- $\sim$ 59 death cases (Guinea)

In March 31, 2014:

- $\sim$ 112 infection cases (Guinea, Liberia, Sierra Leon)
- $\sim$ 70 death cases (Guinea, Liberia, Sierra Leon)

One week later:

- $+39$ new infection cases in W.A.
- $+25$ new death cases in W.A.

Early Detection and Accurate Prediction of Outbreak's Magnitude

At end of July of 2014:

- $\sim$ 1,201 infection cases
- $\sim$ 672 death cases

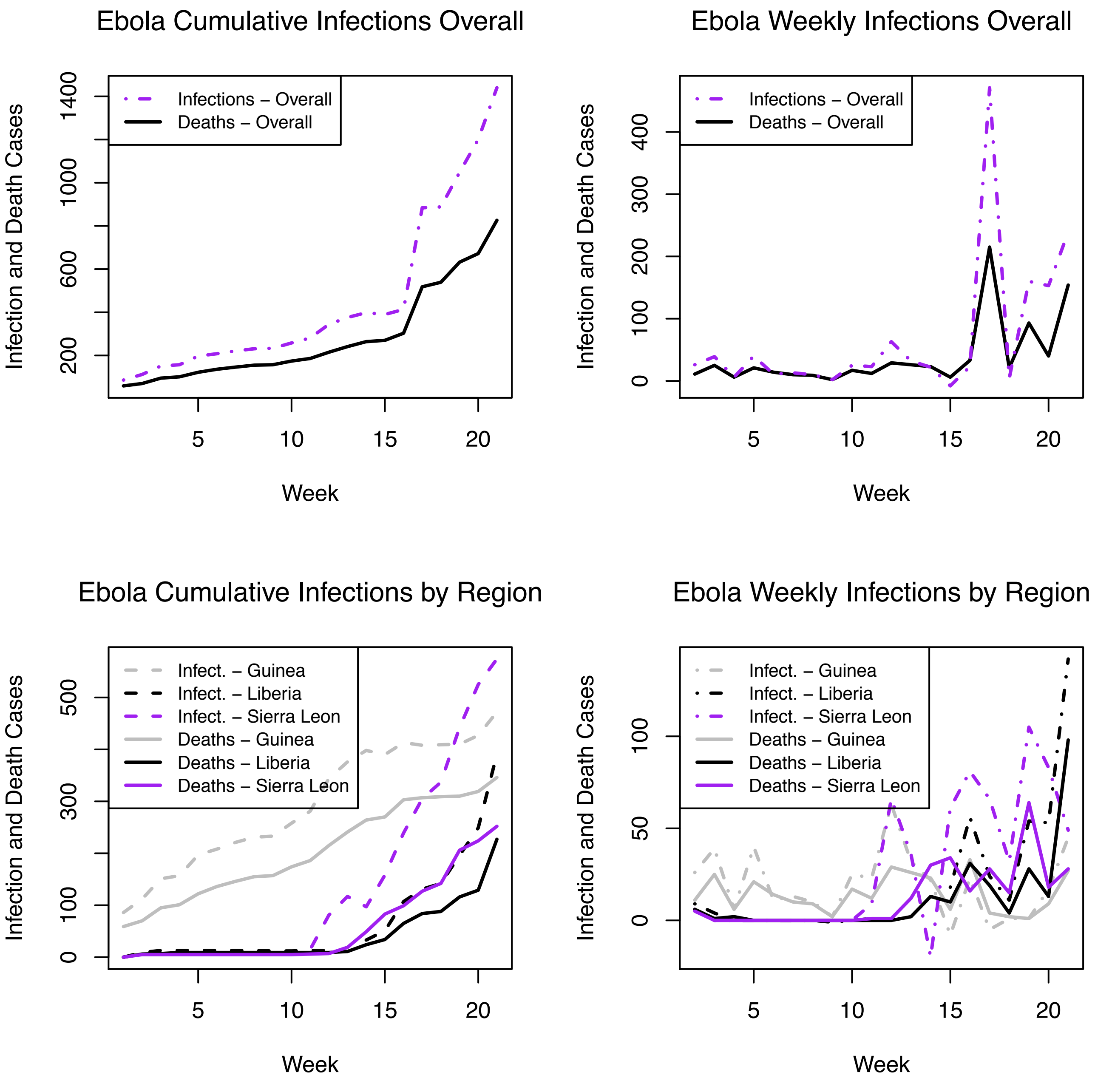
At end of October of 2014:

- $\sim$ 13,268 infection cases
- $\sim$ 4,922 death cases

- This begs the question: Why did the needed interventions not happen earlier? In June, July, or August of 2014?

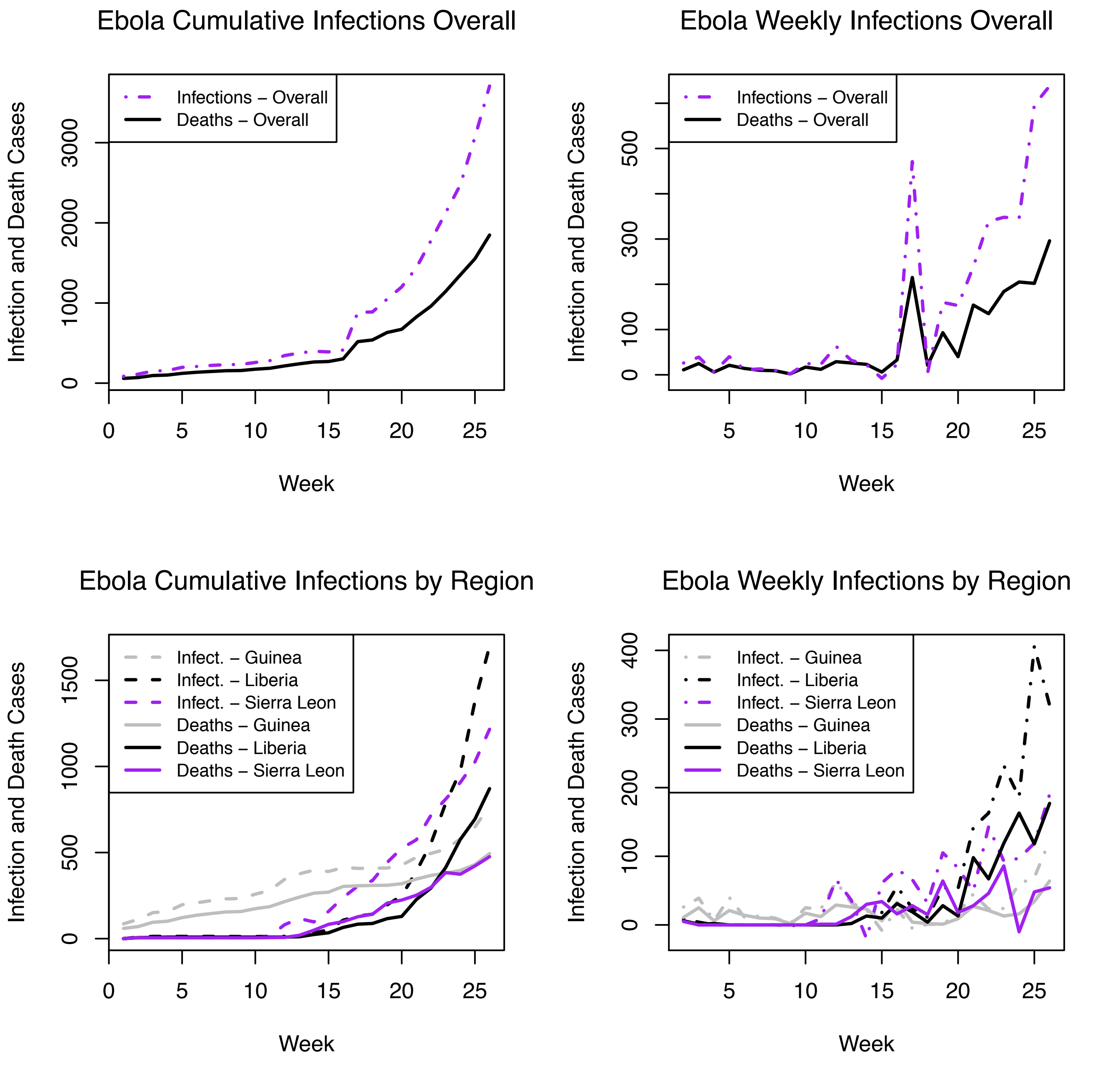
Number of Infected and Death Cases in July 2014

...in July 2014 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
200
600
1000
1400
5
10
15
20
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
100
200
300
400
5
10
15
20
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
100
300
500
5
10
15
20
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
50
100
5
10
15
20
Week

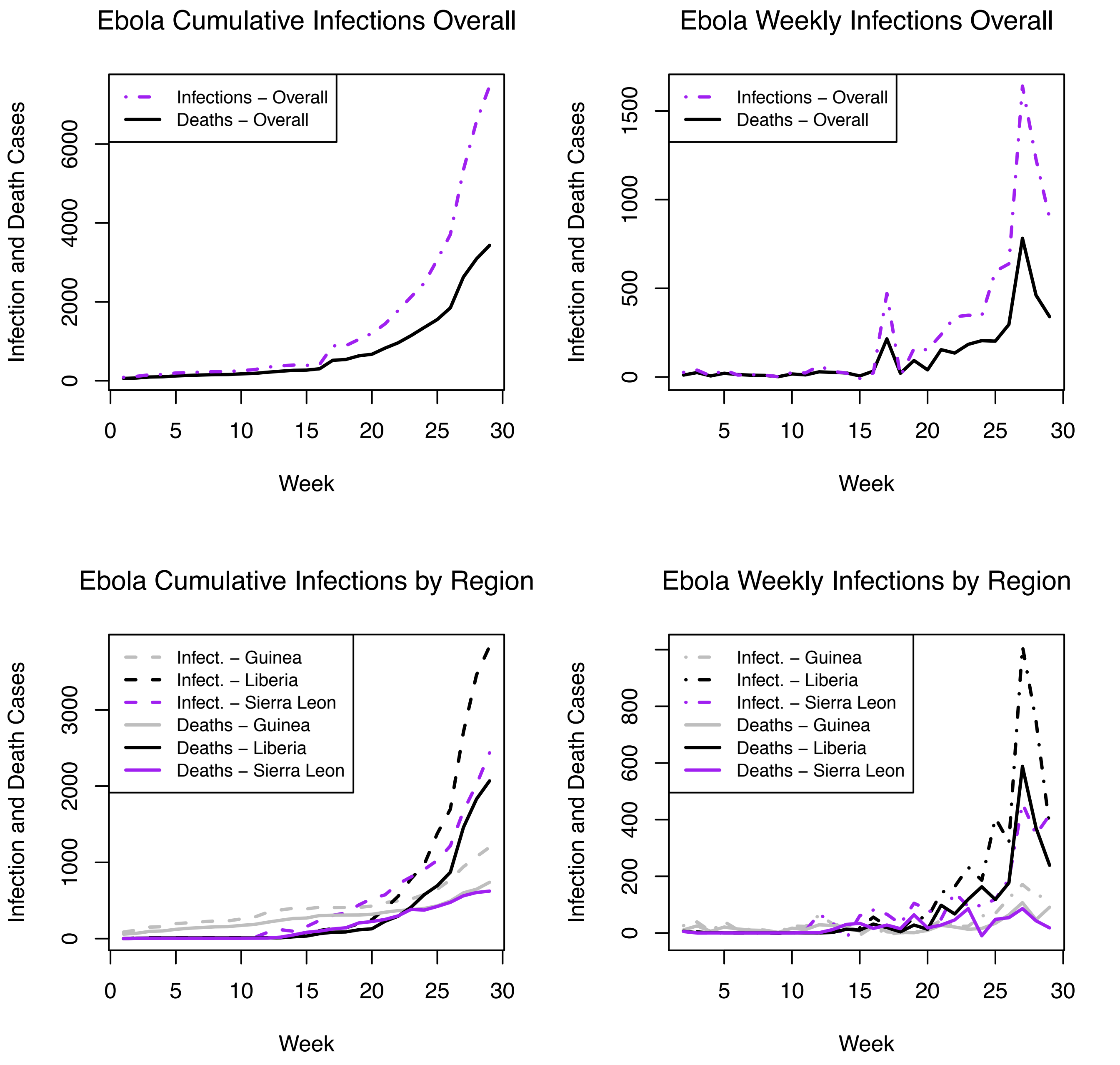
Number of Infected and Death Cases in August 2014

...in August 2014 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
Week
0
1000
2000
3000
0
5
10
15
20
25
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
Week
0
100
300
500
5
10
15
20
25
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
Week
0
500
1000
1500
0
5
10
15
20
25
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
Week
0
100
200
300
400
5
10
15
20
25

Number of Infected and Death Cases in September 2014

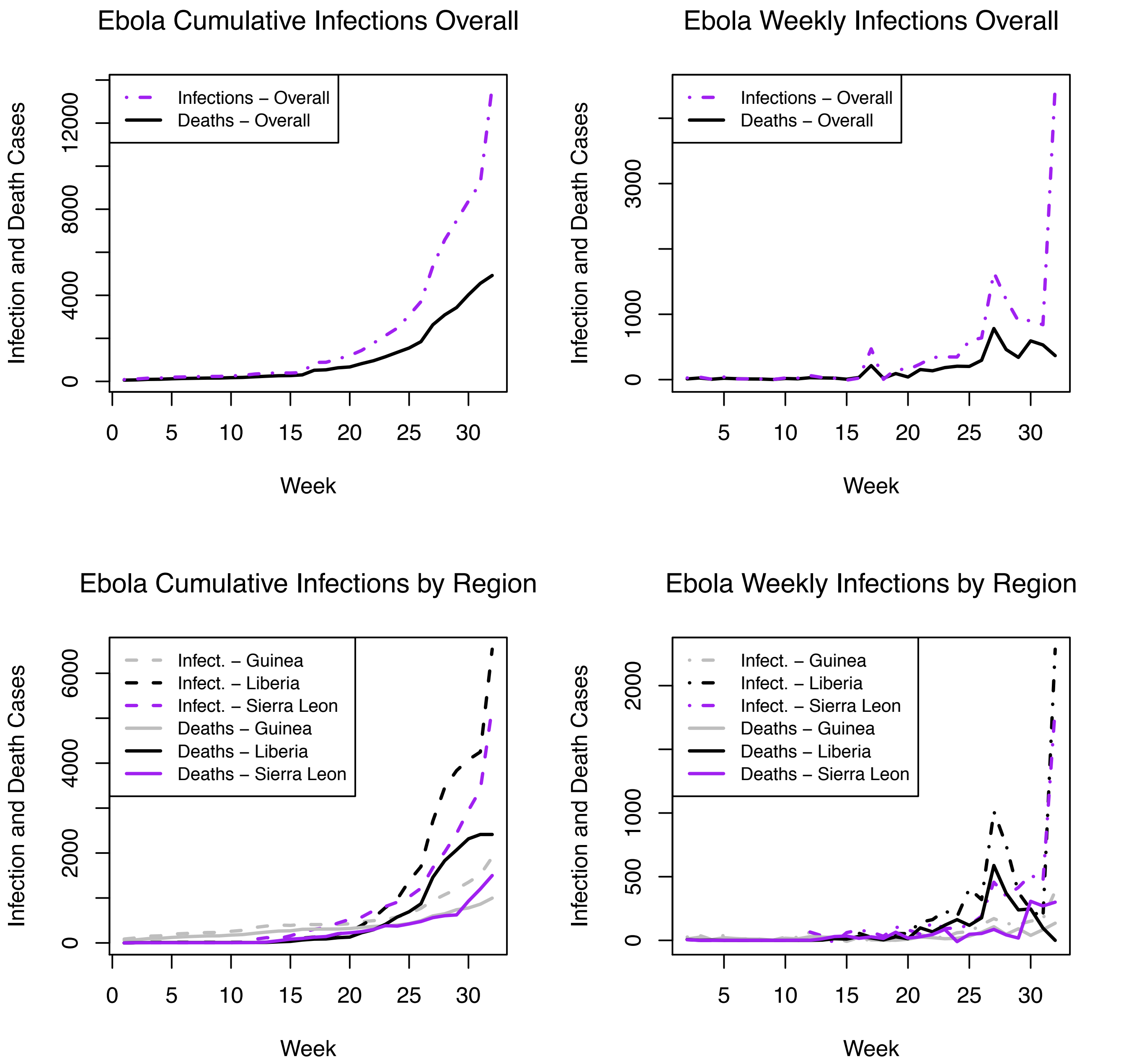
...in September 2014 (end of month)



Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
Week

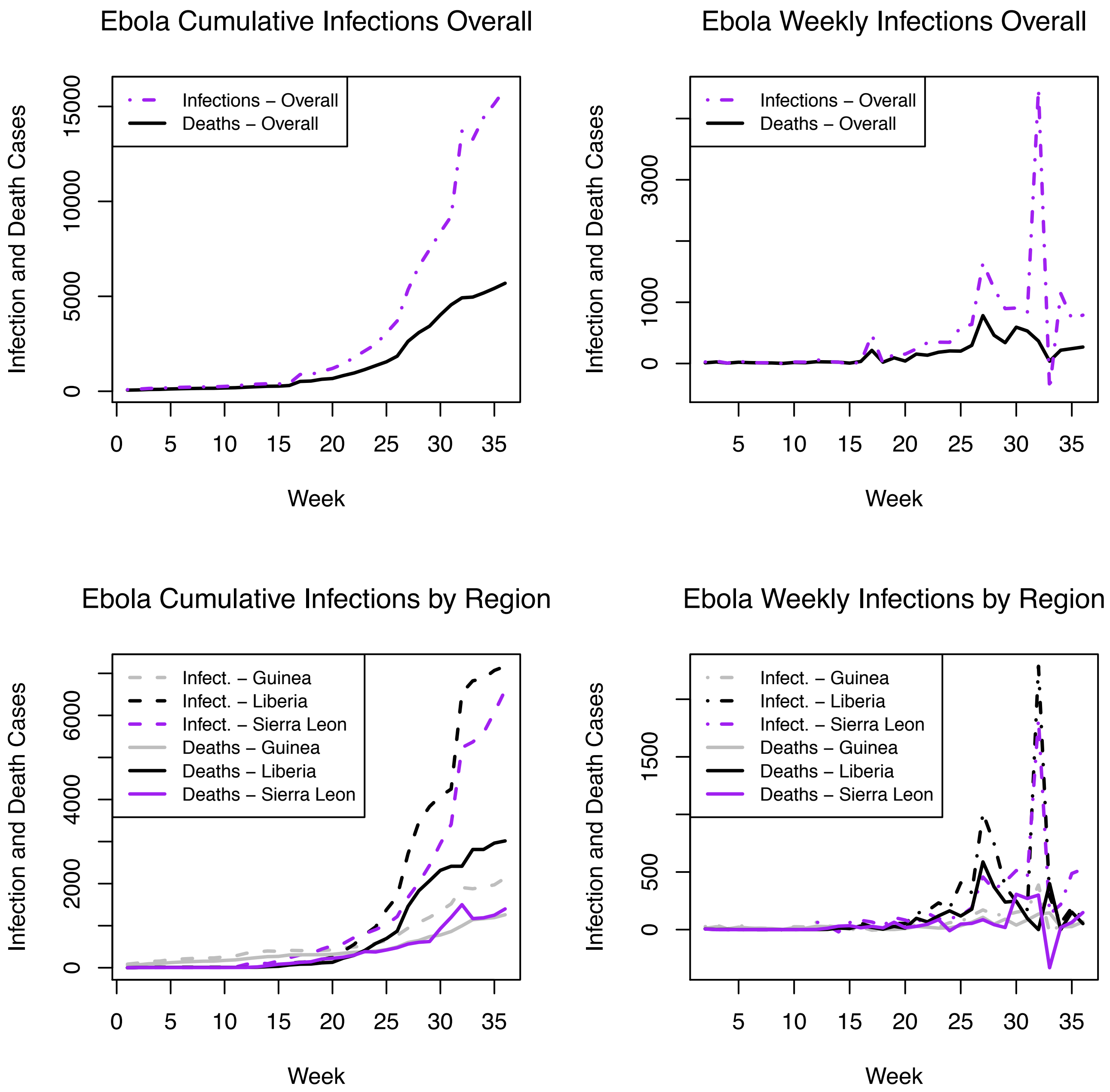
Number of Infected and Death Cases in October 2014

...in October 2014 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
4000
8000
12000
0
5
10
15
20
25
30
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
1000
3000
5
10
15
20
25
30
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
2000
4000
6000
0
5
10
15
20
25
30
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1000
2000
5
10
15
20
25
30
Week

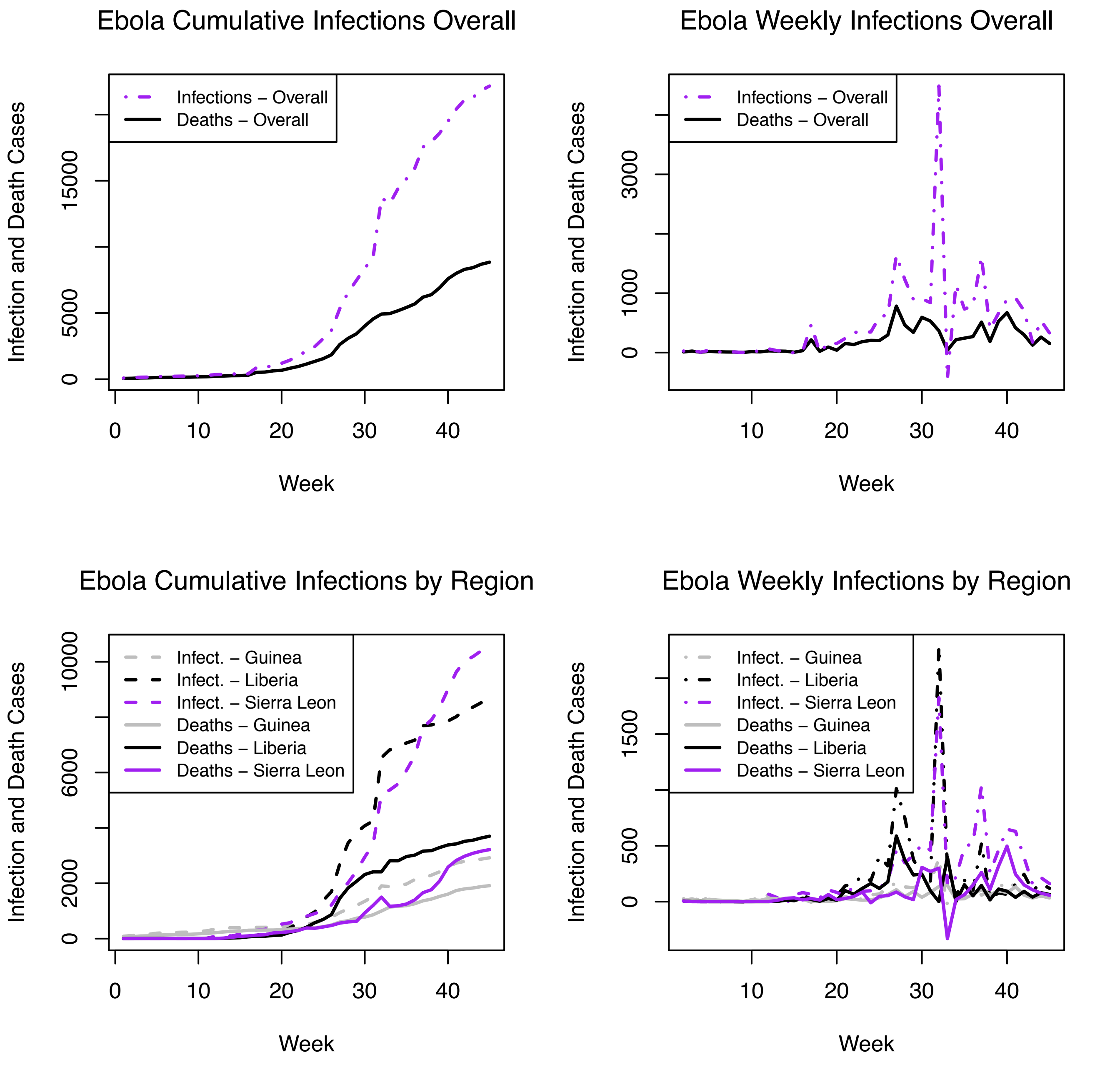
Number of Infected and Death Cases in November 2014

...in November 2014 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
5000
10000
15000
0
5
10
15
20
25
30
35
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
1000
3000
5
10
15
20
25
30
35
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
2000
4000
6000
0
5
10
15
20
25
30
35
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1500
5
10
15
20
25
30
35
Week

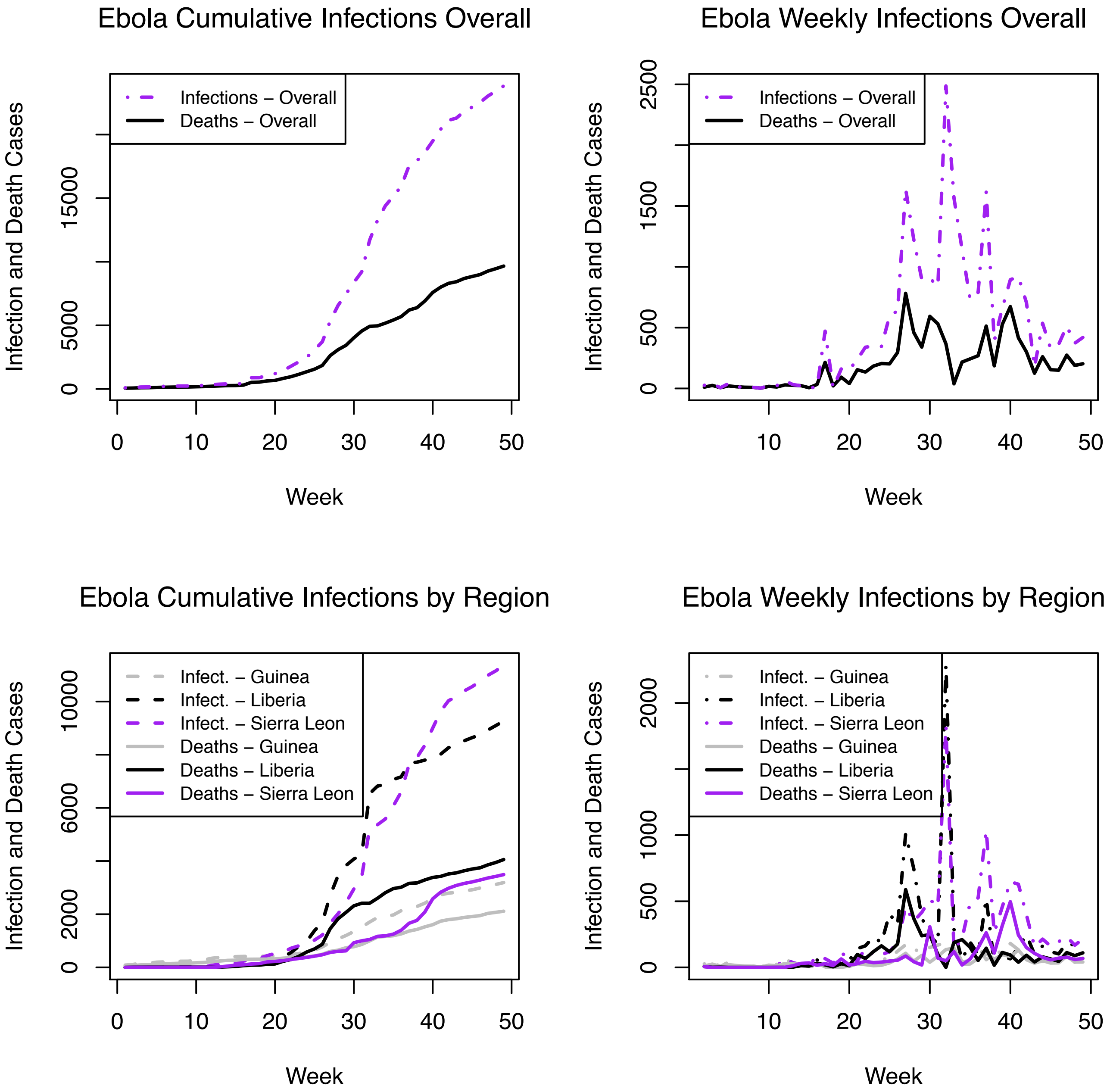
Number of Infected and Death Cases in January 2015

...in January 2015 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
5000
15000
0
10
20
30
40
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
1000
3000
10
20
30
40
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
2000
6000
10000
0
10
20
30
40
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1500
10
20
30
40
Week

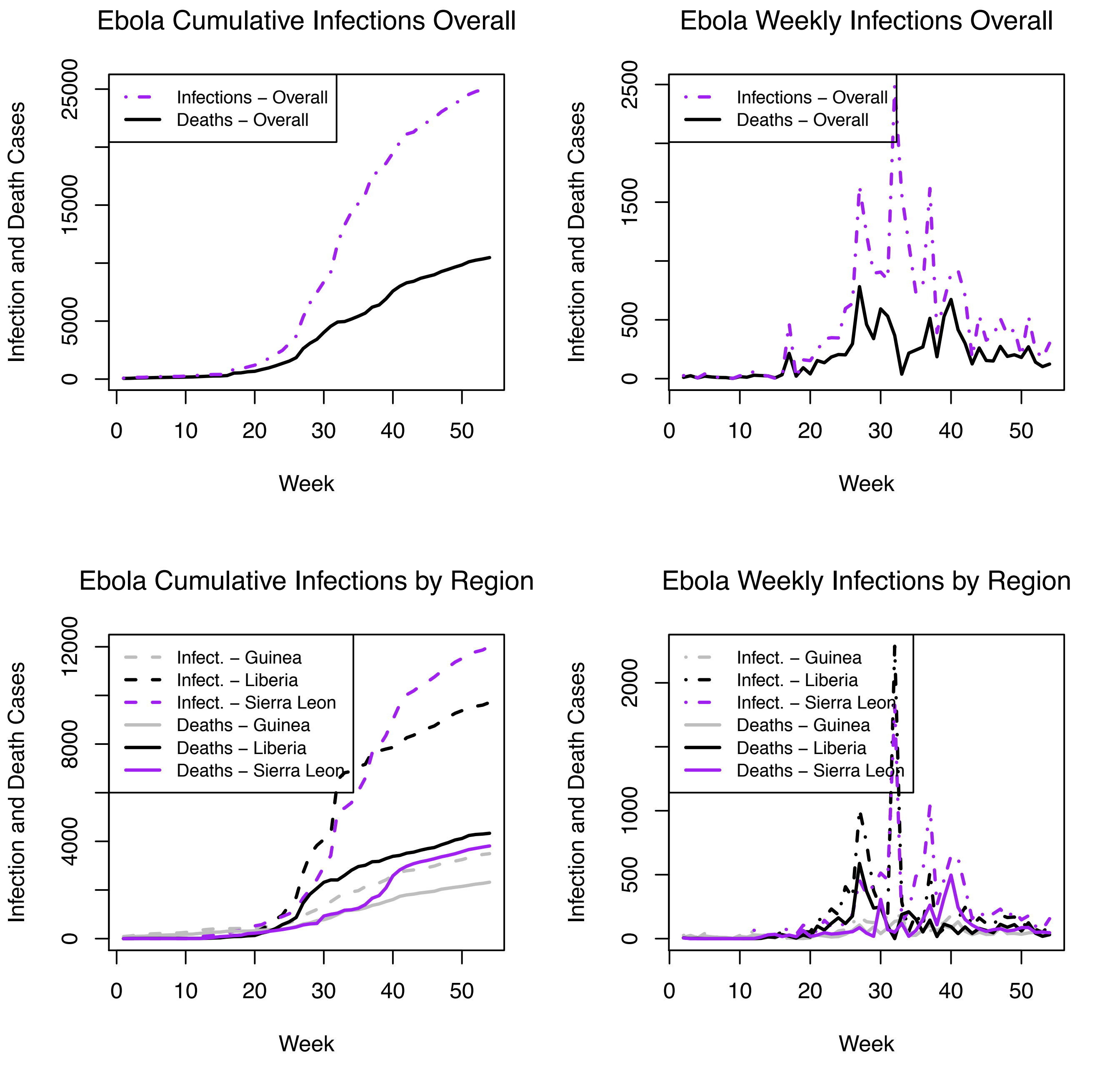
Number of Infected and Death Cases in February 2015

...in February 2015 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
5000
15000
0
10
20
30
40
50
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
500
1500
2500
10
20
30
40
50
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
2000
6000
10000
0
10
20
30
40
50
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1000
2000
10
20
30
40
50
Week

Number of Infected and Death Cases in March 2015

...in March 2015 (end of month)


Ebola Cumulative Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
5000
15000
25000
0
10
20
30
40
50
Week
Ebola Weekly Infections Overall
Infections – Overall
Deaths – Overall
Infection and Death Cases
0
500
1500
2500
0
10
20
30
40
50
Week
Ebola Cumulative Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
4000
8000
12000
0
10
20
30
40
50
Week
Ebola Weekly Infections by Region
Infect. – Guinea
Infect. – Liberia
Infect. – Sierra Leon
Deaths – Guinea
Deaths – Liberia
Deaths – Sierra Leon
Infection and Death Cases
0
500
1000
2000
0
10
20
30
40
50
Week

Estimation of Intervention Effect During Outbreak

In late January of 2015:

- $\sim$ 21,832 infection cases
- $\sim$ 8,690 death cases
- $\sim$ 600 new infection cases per week in combined W.A. region
- $\sim$ 300 new death cases per week in combined W.A. region.
- Was first time back to levels last seen in August of 2014, while weekly numbers now were decreasing.

Outbreak after Intervention

At end of October of 2014 - during intervention period

- $\sim$ 13,268 infection cases
- $\sim$ 4,922 death cases

In late January of 2015 - soon after interventions

- $\sim$ 21,832 infection cases
- $\sim$ 8,690 death cases

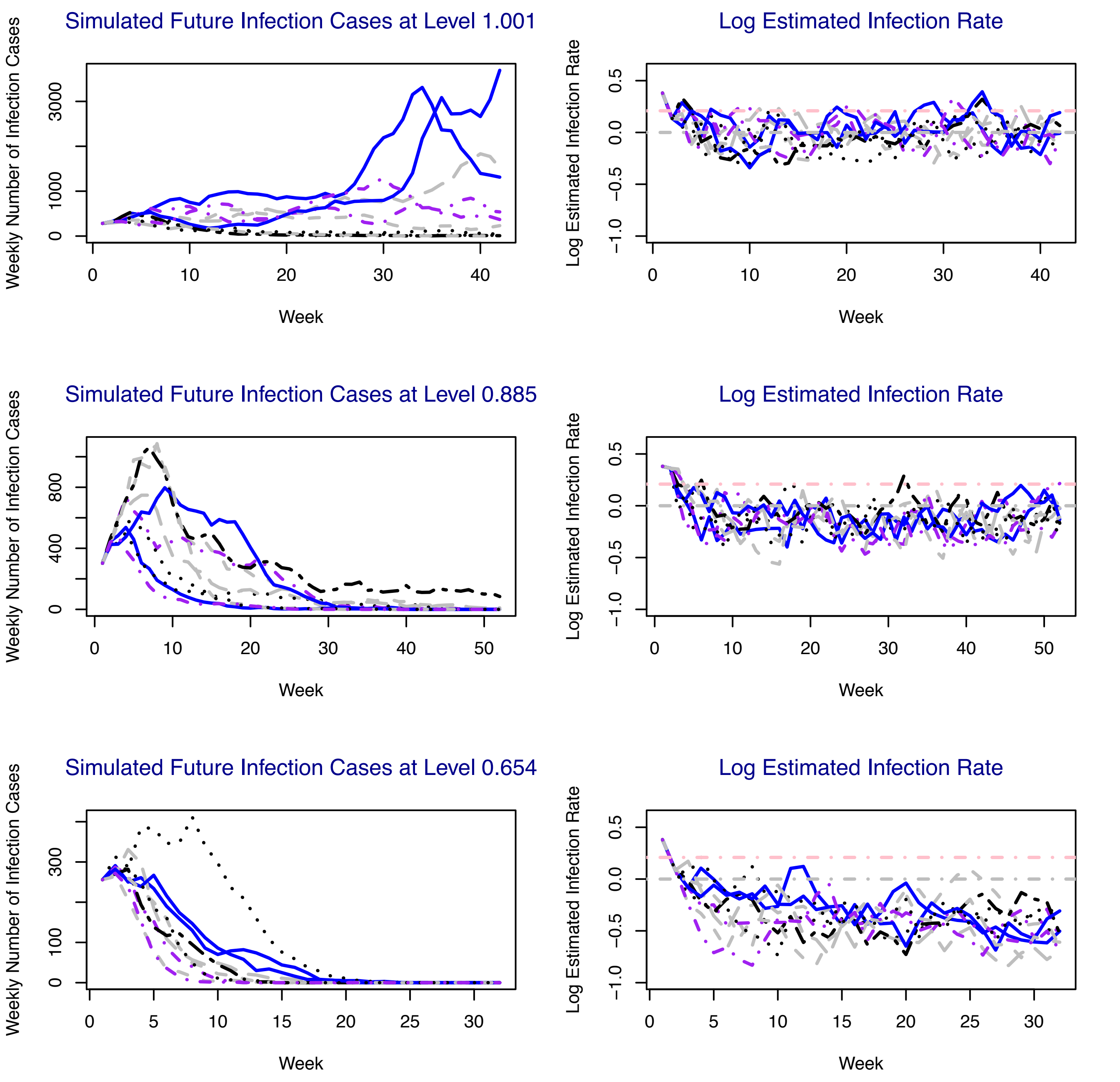
At end of September of 2015 - close to end of outbreak

- $\sim$ 28,319 infection cases
- $\sim$ 11,296 death cases

Impact of Intervention

A Statistical Model for Virus Outbreak Should Help to Answer the Following Questions:

- Does a certain intervention have a sufficient impact?
- Does intervention stabilize outbreak?
- Will intervention eventually end outbreak?
- Expected time to halt outbreak when there are zero or few infection cases?


Simulated Future Infection Cases at Level 1.001
Weekly Number of Infection Cases
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
Week
Simulated Future Infection Cases at Level 0.885
Weekly Number of Infection Cases
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
Week
Simulated Future Infection Cases at Level 0.654
Weekly Number of Infection Cases
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
Week

Mathematical Model for Virus or Infection Outbreaks

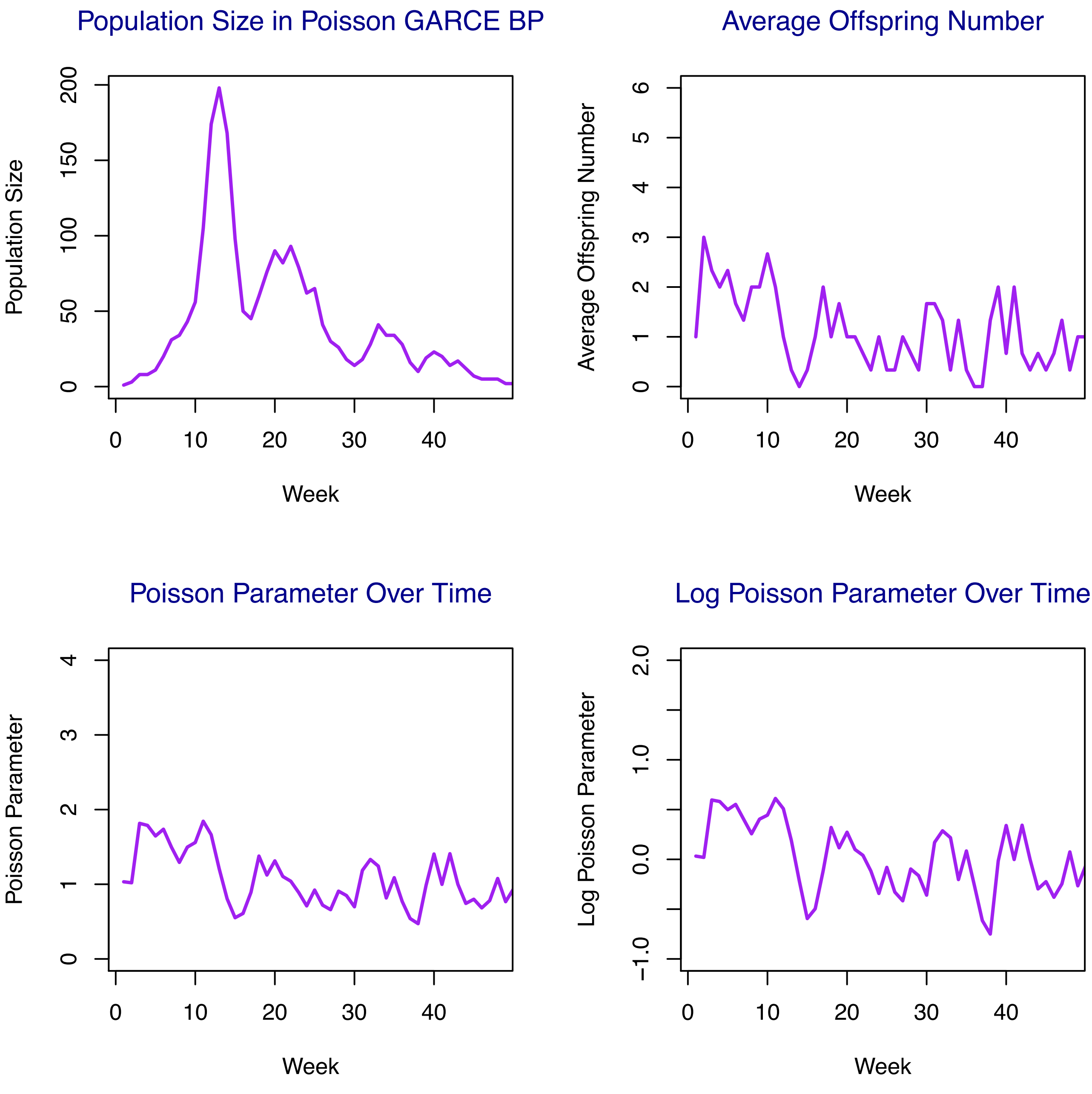
- This example illustrates that outbreaks of infections, avian flus, and viruses spread, grow, or change at rates that vary with time – at peak rates during pandemic time periods, while at low rates when near extinction or in remission.
- A widely studied mathematical model that has been applied to spatially model spread of epidemics, infectious diseases, cancerous tumor growth, and social network traffic consists in the branching processes in random environments (BPRE).

Mathematical Model for Virus or Infection Outbreaks

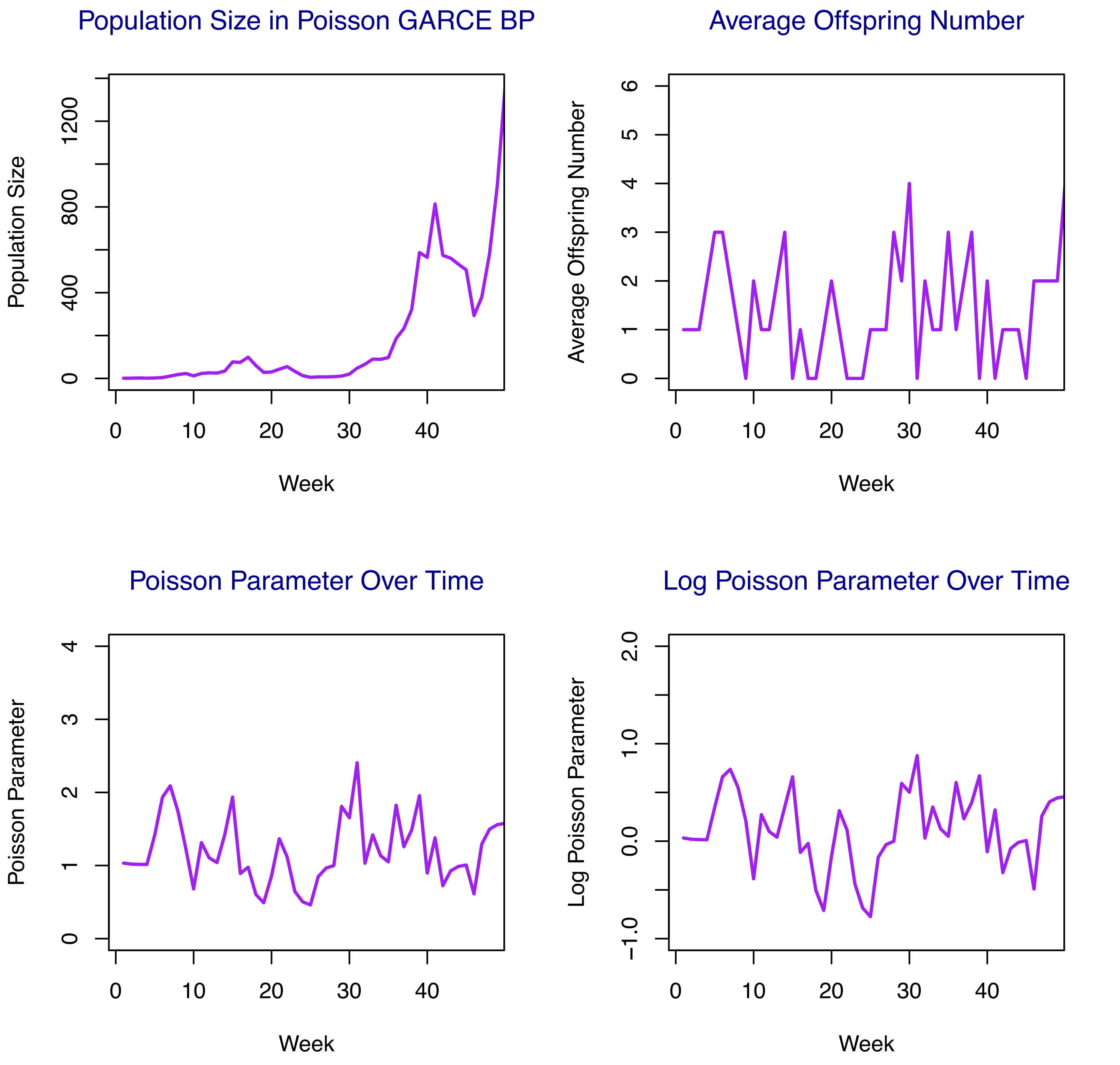
- Yet the BPREs do not allow for such time-varying and dynamic environments.
- We will propose a novel and simple approach to BPREs whose environments follow a time series model and are dynamic.
- These will allow for time-varying random environments and instances of peak growth and near extinction-type rates.
- While residing at the interface of time series (TS) and branching processes, they can be analyzed via TS techniques.
- Maximum likelihood estimation (MLE) for model building, followed by forecasting approaches of future values or events of the process are readily available and implementable.

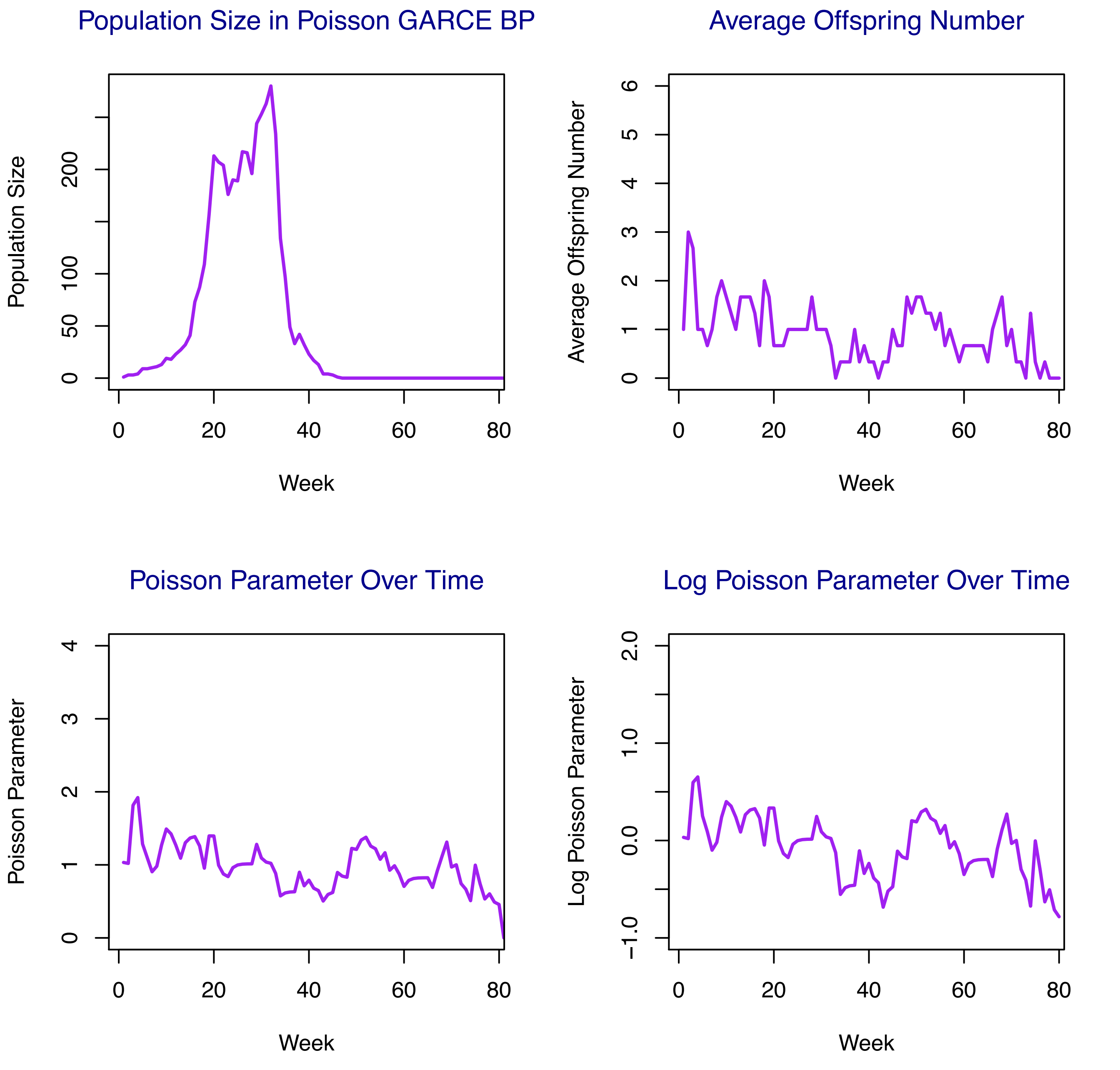
Simulations of Poisson GARCE Branching Process

Look at a few simulations to mimic possible outbreak trajectories

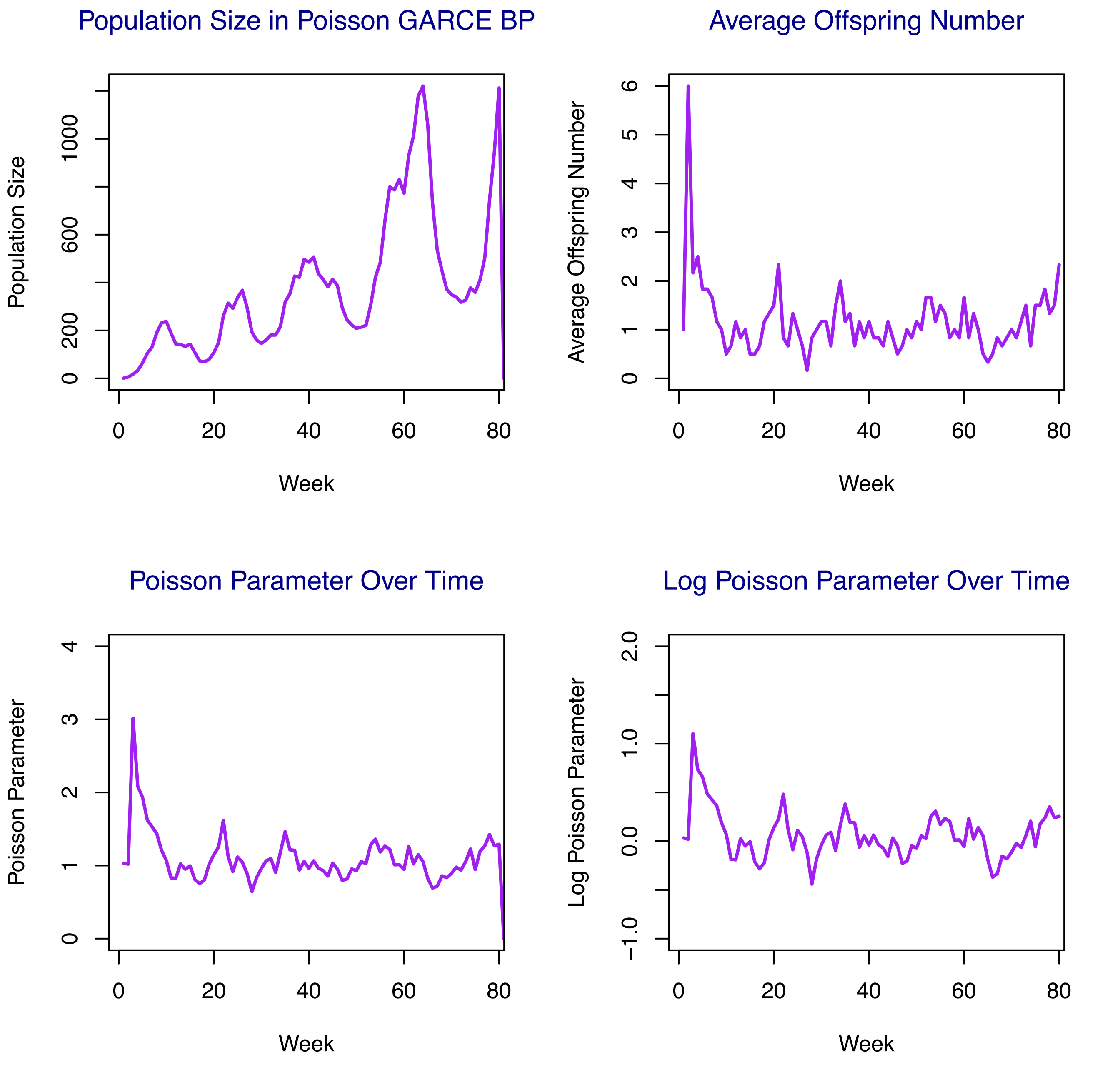


Population Size in Poisson GARCE BP
Population Size
0
50
100
150
200
0
10
20
30
40
Week
Average Offspring Number
Average Offspring Number
0
1
2
3
4
5
6
0
10
20
30
40
Week
Poisson Parameter Over Time
Poisson Parameter
0
1
2
3
4
0
10
20
30
40
Week
Log Poisson Parameter Over Time
Log Poisson Parameter
−1.0
0.0
1.0
2.0
0
10
20
30
40
Week


Population Size in Poisson GARCE BP
Population Size
0
400
800
1200
Week
0
10
20
30
40
Average Offspring Number
Average Offspring Number
0
1
2
3
4
5
6
Week
0
10
20
30
40
Poisson Parameter Over Time
Poisson Parameter
0
1
2
3
4
Week
0
10
20
30
40
Log Poisson Parameter Over Time
Log Poisson Parameter
−1.0
0.0
1.0
2.0
Week
0
10
20
30
40



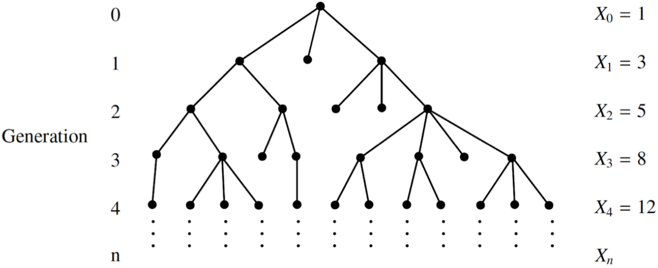
Population Size in Poisson GARCE BP
Population Size
0
50
100
200
Week
0
20
40
60
80
Average Offspring Number
Average Offspring Number
0
1
2
3
4
5
6
Week
0
20
40
60
80
Poisson Parameter Over Time
Poisson Parameter
0
1
2
3
4
Week
0
20
40
60
80
Log Poisson Parameter Over Time
Log Poisson Parameter
−1.0
0.0
1.0
2.0
Week
0
20
40
60
80


Population Size in Poisson GARCE BP
Population Size
0
200
600
1000
0
20
40
60
80
Week
Average Offspring Number
Average Offspring Number
0
1
2
3
4
5
6
0
20
40
60
80
Week
Poisson Parameter Over Time
Poisson Parameter
0
1
2
3
4
0
20
40
60
80
Week
Log Poisson Parameter Over Time
Log Poisson Parameter
−1.0
0.0
1.0
2.0
0
20
40
60
80
Week

Branching Processes in Random Environments

Branching Process (BP)

is represented by a population size $\{Z_t\}_{t \geq 0}$ at time t with initial size $Z_0 > 0$, where each of the Z_t members reproduces offspring according to a common offspring distribution with probability generating function (p.g.f.) $\varphi(s)$.


Generation
0
1
2
3
4
n
$X_0 = 1$
$X_1 = 3$
$X_2 = 5$
$X_3 = 8$
$X_4 = 12$
X_n

Branching Processes in Random Environments

Branching Process in Random Environments (BPRE)

A branching process $\{Z_t\}_{t \geq 0}$ with offspring distribution $\sim$ p.g.f. $\varphi_{\lambda_t}(s)$, where parameter λ_t is a *random* variable.

Note: Research on these goes back to

- Smith and Wilkinson,1968–1971
- Athreya and Karlin, 1971.

Smith-Wilkinson Model, 1969-1971

When $\varphi_{\lambda_t}(s)$ is assumed to be chosen *independently* at random from a collection of p.g.f.s with specified time-homogeneous distribution.

Branching Processes in Random Environments

BPRE in Dynamic Environment

- Define a BPRE with dynamic environments, where sampling distribution for $\varphi_{\lambda_t}(s)$ evolves dynamically at any time.
- Parameter λ_t is governed by a recurrence relation of the form

$$\lambda_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, Y_{t-i} + \sum_{j=1}^{q} \beta_j \, \lambda_{t-j},$$

  where $\{Y_t\}$ is an estimate for mean offspring number of previous generation and the α_i and β_j are model parameters.
- Thus, λ_t is regressed on p past values of internal process parameter and q past values of Y_t. Choose p and q suitably.

Branching Processes in Random Environments

Branching Process in Generalized Autoregressive Conditional Environments (GARCE BP)

Let us define a

- branching process in **"Generalized AutoRegressive Conditional Environments"** or
- **GARCE branching process.**

Here is notation to get started...

GARCE Branching Processes

Notation:

- Let $Z_{t+1} = \sum_{i=1}^{Z_t} X_{t,i}$.

  Thus, the $\{X_{t,k}\}_{1 \leq k \leq Z_t}$ are t-th generation offspring numbers of Z_t particles of generation t.

- Let $Y_{t+1} = \frac{1}{K} \sum_{k=1}^{K} X_{t,k}$ for fixed small integer $K \geq 1$.

  So, Y_{t+1} estimates the mean offspring number per parent in generation t.

- Environment parameter $\lambda = \{\lambda_t(\omega)\}_{t \geq 0} = (\lambda_0, \lambda_1, \lambda_2, \ldots)$. Represents infection rate of the virus.

GARCE Branching Processes

Notation (cont'd):

- $\mathcal{M}$: Collection of probability distributions

$$\{\{p_i\}_{i=0}^{\infty}, \sum_{i\geq 1} i \cdot p_i < \infty, \ 0 \leq p_0 + p_1 < 1\}$$

- $\mathcal{B}$: Borel σ-algebra in $\mathcal{M}$
- Seq. of mappings $\{\lambda_t(\omega)\}_{t\geq 0}$ from $(\Omega, \mathcal{F}, \mathbf{P})$ into $(\mathcal{M}, \mathcal{B})$
- For any such $\lambda = \lambda(\omega) = \{\lambda_t(\omega)\}_{t\geq 0}$, define p.g.f.

$$\varphi_\lambda(s) \;=\; \sum_{i=0}^{\infty} p_i(\lambda)\, s^i \qquad\qquad |s| \leq 1.$$

GARCE Branching Processes

Notation (cont'd):

- $\sigma(D)$: sub-σ-algebra of $\mathcal{F}$ generated by collection D of random variables (r.v.s) on $(\Omega, \mathcal{F}, \mathbf{P})$
- σ-algebras

$$\begin{aligned} \mathcal{F}_t(\lambda) &= \sigma(\lambda_0, \lambda_1, \ldots, \lambda_t), \quad \mathcal{F}(\lambda) = \sigma(\lambda) \\ \mathcal{F}_{t,z,y}(\lambda) &= \sigma(\lambda_0, \lambda_1, \ldots, \lambda_t, Z_0, Z_1, \ldots, Z_t, Y_0, Y_1, \ldots, Y_t). \end{aligned}$$

Represents historical information known of the process or virus.

GARCE Branching Processes

BP in Generalized AutoRegressive Conditional Environments

A GARCE(p, q) branching process of order p and q is a process $\{Z_t\}_{t\geq 0}$ with environmental process $\{\lambda_t\}_{t\geq 0}$ that satisfies

$$\begin{aligned} \mathbf{E}(s^{Z_{t+1}} \mid \mathcal{F}_{t,z,y}) &= [\varphi_{\lambda_{t+1}}(s)]^{Z_t} \quad \text{a.s.} \\ \lambda_{t+1} &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \, Y_{t+1-i} + \sum_{j=1}^{q} \beta_j \, \lambda_{t+1-j} \\ &\quad (\alpha_0 > 0, \alpha_i, \beta_j \geq 0, \alpha_p, \beta_q > 0, p \geq 1, q \geq 0), \end{aligned}$$

and almost surely,

$$\mathbf{E}(s_1^{Z_{t_1}} \cdot \ldots \cdot s_k^{Z_{t_k}} | \mathcal{F}(\lambda), Z_0 = m) = [\mathbf{E}(s_1^{Z_{t_1}} \cdot \ldots \cdot s_k^{Z_{t_k}} | \mathcal{F}(\lambda), Z_0 = 1)]^m.$$

Remarks: GARCE Branching Processes

Remarks:

- First relation expresses that
  (a) $X_{t,k}$ are identically distributed $\sim \varphi_{\lambda_{t+1}}(s)$ and
  (b) $X_{t,k}$ are conditionally independent, given $\mathcal{F}_{t,z,y}$.
- Third relation assures that each child's branching process generated has the same distribution as its parent's branching process created from first generation onwards.
- Note that only the second relation is not part of the definition of the usual BPRE and is new.

Remarks: GARCE Branching Processes

Remarks (cont'd):

- Starting values set: $\lambda_t = \lambda_*$ (some λ_*) > 0, $Y_t = 0$ for all $t \leq 0$, $Z_0 = 1$, and $Y_1 = Z_1$.
- Existence of such a process $\{Z_t\}_{t \geq 0}$ is assured by Harris construction, 1963.
- Key Structural Property: Conditionally on $\{\lambda_t\}_{t \geq 0}$, process $\{Z_t\}_{t \geq 0}$ is Markovian with independent lines of descent.

Observations: GARCE Branching Processes

Observations:

- Importantly, GARCE BP model allows for application of time series techniques in context of branching processes.
- It is a BPRE whose environmental and mean offspring number processes exhibit autoregressive serial dependence structure. Allows for non-linear behavior and clustering of outliers.

Example Offspring Distributions

Poisson INGARCH Offspring

- $X_{t,k} \,|\, \mathcal{F}_{t,z,y} \sim \mathcal{P}(\lambda_{t+1})$ for fixed $t \geq 1$ and each $1 \leq k \leq Z_t$
  $\mathcal{P}(\lambda_{t+1})$: Poisson distribution w. parameter λ_{t+1}.

Negative Binomial INGARCH Offspring

- $X_{t,k} \,|\, \mathcal{F}_{t,z,y} \sim \mathcal{NB}(r, p_{t+1})$ for each $1 \leq k \leq Z_t$
  $\mathcal{NB}(r, p_{t+1})$: negative binomial distr. w. parameters r (integer $r \geq 1$), $p_{t+1} \in (0,1)$ with $(1 - p_{t+1})/p_{t+1} = \lambda_{t+1}$.

- Further distributions: generalized Poisson and binomial.
- For illustration purposes, restrict discussion to Poisson case.

Stationarity

Theorem

Consider Poisson case and suppose $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1.$

- *There exists a unique strictly stationary mean offspring number process* $\{Y_t\}_{t \in \mathbf{Z}}$ *for the GARCE BP prior to the random extinction event* $Z_T = 0$ *(if any).*
- *First two moments of* $\{Y_t\}_t$ *are finite and expressions are*

$$\begin{aligned} \mathbf{E}(Y_t) &= \mathbf{E}(\lambda_t) = \mu := \alpha_0/(1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j), \\ \mathrm{Var}(Y_t) &= \mathrm{Var}(\lambda_t) + \mu/K. \end{aligned}$$

Auto- and Cross-Correlations

For stationary processes $\{Y_t\}_t$ and $\{\lambda_t\}_t$, denote their respective autocovariance function (ACVF)

- $\{\gamma_Y(k) = \mathrm{Cov}(Y_{t+k}, Y_t)\}_{k\geq 0}$
- $\{\gamma_\lambda(k) = \mathrm{Cov}(\lambda_{t+k}, \lambda_t)\}_{k\geq 0}$

and their cross-covariance function (CVF)

- $\{\gamma_{Y\lambda}(k) = \mathrm{Cov}(Y_{t+k}, \lambda_t)\}_{k\geq 0}$.

Auto- and Cross-Correlations

Theorem

Consider Poisson case and suppose $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$.
ACVFs $\{\gamma_Y(k)\}_{k\geq 0}$ *and* $\{\gamma_\lambda(k)\}_{k\geq 0}$ *obey two intralinked linear recurrence relations. Cross-CVF* $\{\gamma_{Y\lambda}(k)\}_{k\geq 0}$ *is given by*

$$\gamma_{Y\lambda}(k) = \begin{cases} \gamma_\lambda(k), & \text{for } k \geq 0 \\ \gamma_Y(k), & \text{for } k < 0. \end{cases}$$

Poisson GARCE(1,1) Branching Process

Poisson GARCE(1, 1) BP with $\alpha_1 + \beta_1 < 1$

Variances and auto-correlation functions (ACFs) of Y_t and λ_t are given by

$$\mathrm{Var}(Y_t) = \frac{\mu}{K} \cdot \frac{1 - \beta_1(2\alpha_1 + \beta_1)}{1 - (\alpha_1 + \beta_1)^2}$$

$$\mathrm{Var}(\lambda_t) = \frac{\mu}{K} \cdot \frac{\alpha_1^2}{1 - (\alpha_1 + \beta_1)^2}$$

$$\rho_Y(k) = (\alpha_1 + \beta_1)^{k-1}\,\alpha_1 \cdot \frac{1 - \beta_1(\alpha_1 + \beta_1)}{1 - \beta_1(2\alpha_1 + \beta_1)}$$

$$\rho_\lambda(k) = (\alpha_1 + \beta_1)^k.$$

Ergodicity for Poisson GARCE(1,1) BP

- Study of survival behavior of the GARCE BP and properties of normalized process are manageable under assumptions of stationarity and ergodicity of bivariate process $\{(Y_t, \lambda_t)\}_t$.
- Ergodicity feature is also crucial to asymptotic theory of conditional maximum likelihood estimators (MLE) in GARCE BP model.

Theorem

If $\alpha_1 + \beta_1 < 1$, the process $\{(Y_t, \lambda_t)\}_{t \in \mathbf{Z}}$ has a unique stationary distribution and $\{Y_t\}_{t \in \mathbf{Z}}$ and $\{(Y_t, \lambda_t)\}_{t \in \mathbf{Z}}$ are ergodic.

MLE for Poisson GARCE(1,1) BP

Theorem

Suppose $\alpha_1 + \beta_1 < 1.$

Let $\theta = (\alpha_0, \alpha_1, \beta_1)'$ *and* θ° *denote true unknown value of* θ.

Under mild assumptions, MLE $\hat{\theta}$ *is unique, consistent, and asymptotically normal*

$$\sqrt{n}(\hat{\theta} - \theta^\circ) \xrightarrow{d} \mathcal{N}(0, G^{-1})$$

for some computable matrix G.

MLE for Poisson GARCE(1,1) BP

- Estimation of model parameters is part of model building that is used to generate forecasts for future values of process.
- First produce predicted values of the environmental process $\{\lambda_t\}$ that are obtained from GARCE recurrence relation.
- Second apply bootstrapping techniques in order to predict future values for processes $\{X_{t,i}\}$ and $\{Z_t\}$.
- Computations for MLE approach are implementable in R.

Certain and Non-Certain Extinction

Results Established for GARCE BP:

- Survival-extinction dichotomy: Certain extinction has probability 0 or 1.
- There are necess. & suff. conditions for non-certain extinction
- Extinction or explosion happens with probability 1
- Characterization of phase transition between subcritical and supercritical GARCE BP
- Survival behavior in these two phases and at criticality.

Extinction Set and Probabilities

Notation:

- Extinction set $\mathcal{E} = \{\omega : Z_t(\omega) = 0 \text{ for some } t\}$
- Conditional extinction probabilities
  $q_k(\lambda) = \mathbf{P}(\mathcal{E} \mid \mathcal{F}(\lambda), Z_0 = k)$
- Unconditional extinction probabilities
  $q_k = \mathbf{P}(\mathcal{E} \mid Z_0 = k)$
- Write $q(\lambda) = q_1(\lambda)$.
- Back-shift transformation:
  $T\lambda = T(\lambda_0, \lambda_1, \ldots) = (\lambda_1, \lambda_2, \ldots)$.

Sequence of Extinction Probabilities

Observe that

$$q_k(\lambda) \;=\; q_1(\lambda)^k \text{ a.s.}, \quad q_k = \mathbf{E}[q_1(\lambda)^k],$$

which establishes that $\{q_k\}_{k\geq 1}$ is a moment sequence.
Furthermore,

$$q(\lambda) = \lim_{t\to\infty} \varphi_{\lambda_1}(\varphi_{\lambda_2}(\ldots\varphi_{\lambda_t}(0)\ldots)),$$

consequently,

$$q(\lambda) \;=\; \varphi_{\lambda_1}(q(T\lambda))$$

It was shown that $q(\lambda)$ is the *minimal* solution, and when $\mathbf{P}(q(\lambda) < 1) = 1$, the *unique* solution to functional equation.

Extinction

When extinction of GARCE BP $\{Z_t\}_t$ happens, it occurs almost surely wrt. probability measure of environmental process.

Theorem

$$\mathbf{P}(q(\lambda) = 1) = 0 \quad \text{or} \quad \mathbf{P}(q(\lambda) = 1) = 1.$$

Note $\mathbf{P}(q(\lambda) < 1) = 1$ is referred to as 'non-certain extinction.'

Sub- and Supercritical Processes

Denote

- $V_\lambda = \log \varphi'_{\lambda_1}(1)$.
  For Poisson offspring: $V_\lambda = \log(\lambda_1)$.

Supercritical, Critical, or Subcritical GARCE BP

GARCE BP $\{Z_t\}$ is *supercritical, critical,* or *subcritical* depending on whether $\mathbf{E}(V_\lambda) = \mathbf{E}[\log(\lambda_1)] > 0, = 0,$ or $< 0,$ respectively.

This provides characterization of phase transition between subcritical and supercritical GARCE BP.

Classification Result

Theorem

(Classification) *Suppose that* $\mathbf{E}(V_\lambda)$ *exists.*

(i) *If* $\mathbf{E}(V_\lambda) < 0$, *then* $\mathbf{P}(q(\lambda) = 1) = 1$.

(ii) *If* $\mathbf{E}(V_\lambda) = 0$, *then either* $\mathbf{P}(q(\lambda) = 1) = 1$ *or* $\mathbf{P}(p_1(\lambda_1) = 1) = 1$.

$\mathbf{P}(p_1(\lambda_1) = 1) = 1$ *implies* $\mathbf{P}(Z_t \equiv 1 \ \forall t \mid Z_0 = 1, \mathcal{F}(\lambda)) = 1$ *wp1.*

(iii) **(Law of Large Numbers)** *If* $\mathbf{E}(V_\lambda) > 0$, *then* $\lim_{t\to\infty} t^{-1} \log Z_t = \mathbf{E}(V_\lambda)$ *a.e. on* $\{\omega : \ Z_t(\omega) \to \infty \ \text{as} \ t \to \infty\}$.

Non-certain extinction is precluded unless $\{Z_t\}$ is supercritical.

Extinction-Explosion Dichotomy

Theorem

(Extinction-Explosion Dichotomy) *Either*

$$\mathbf{P}(Z_t \to 0 \ \text{or} \ Z_t \to \infty \,|\, \mathcal{F}(\lambda)) = 1,$$

independently of Z_0, *or*

$$\mathbf{P}(Z_t \equiv 1 \ \forall t \mid Z_0 = 1, \mathcal{F}(\lambda)) = 1$$

with probability 1.

GARCE BP with Intervention

Intervention Effect Added to Model

- Intervention at time τ with effect size ν transforms $\{\lambda_t\}_{t\geq 0}$
- New environmental process $\{\kappa_t\}_{t\geq 0}$ satisfies recurrence

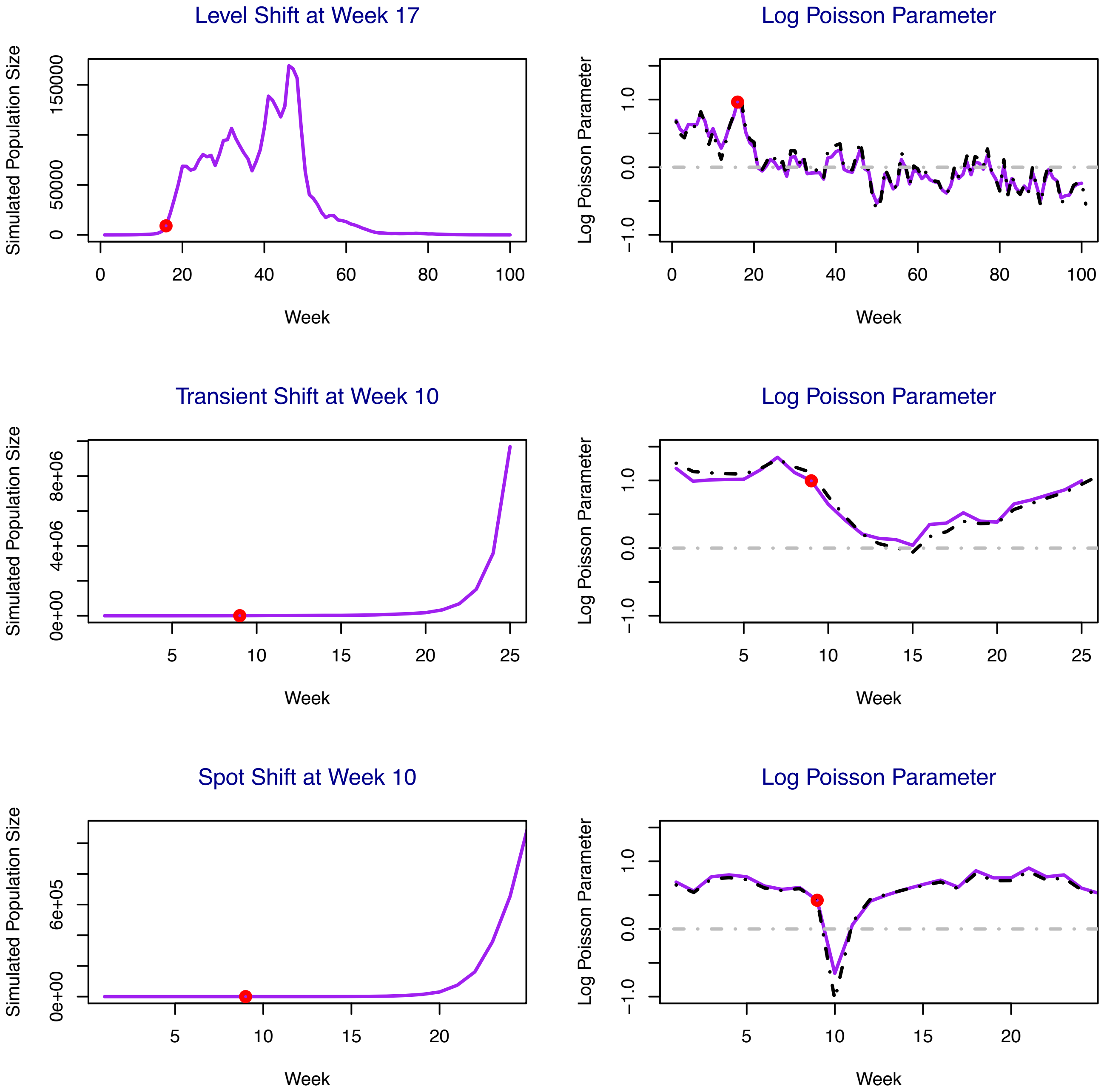
$$\kappa_{t+1} \quad = \quad \alpha_0 + \sum_{i=1}^{p} \alpha_i \, \tilde{Y}_{t+1-i} + \sum_{j=1}^{q} \beta_j \, \kappa_{t+1-j} + \nu \, \zeta_{t+1},$$

where $\zeta_{\tau+h} = \xi_h = \delta^h$ for $h \geq 0$ and $\zeta_t = 0$ for $t < \tau$.
$\delta = 1$ - level shift, $\delta \in (0,1)$ - transient shift, $\delta = 0$ - additive outlier

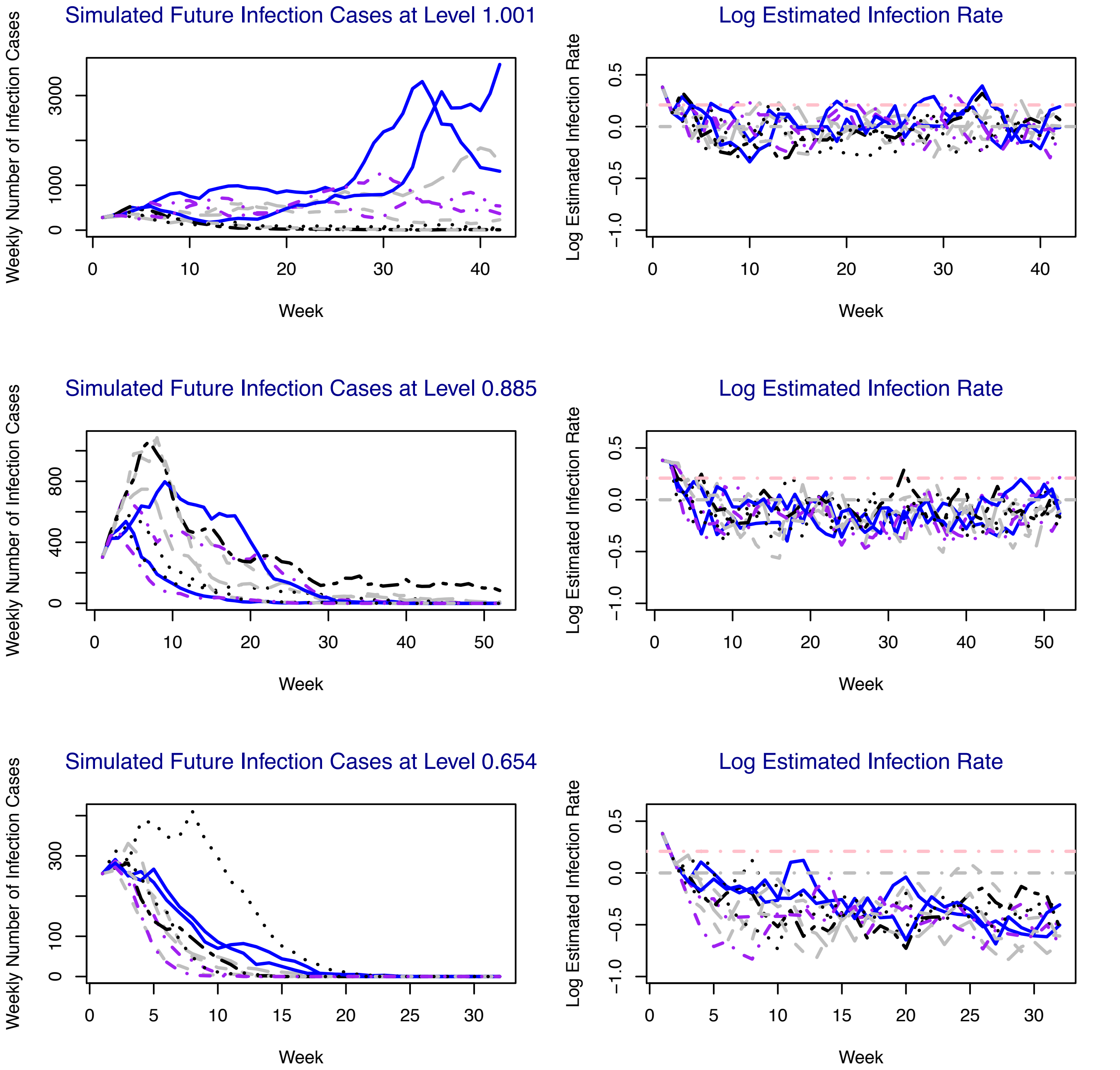
- For $h \geq 0$,
$\kappa_{\tau+h} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, \tilde{Y}_{\tau+h-i} + \sum_{j=1}^{q} \beta_j \, \kappa_{\tau+h-j} + \nu \, \xi_h.$

Simulations

- Simulated Outbreak Trajectories of Infection Cases After 3 Different Types of Interventions
- Simulated Outbreak Paths in Different Regimes After Level Shift Intervention



Level Shift at Week 17
Simulated Population Size
0
50000
150000
0
20
40
60
80
100
Week
Log Poisson Parameter
Log Poisson Parameter
−1.0
0.0
1.0
0
20
40
60
80
100
Week
Transient Shift at Week 10
Simulated Population Size
0e+00
4e+06
8e+06
5
10
15
20
25
Week
Log Poisson Parameter
Log Poisson Parameter
−1.0
0.0
1.0
5
10
15
20
25
Week
Spot Shift at Week 10
Simulated Population Size
0e+00
6e+05
5
10
15
20
Week
Log Poisson Parameter
Log Poisson Parameter
−1.0
0.0
1.0
5
10
15
20
Week



Simulated Future Infection Cases at Level 1.001
Log Estimated Infection Rate
Simulated Future Infection Cases at Level 0.885
Log Estimated Infection Rate
Simulated Future Infection Cases at Level 0.654
Log Estimated Infection Rate
Weekly Number of Infection Cases
Log Estimated Infection Rate
Week

Ebola Data Set Presented

Observations:

- Weekly numbers of infections over time overall & by region
- Time period: March 24, 2014 through February 25, 2015
- Data Source: website www.cdc.gov, Centers for Disease Control and Prevention
- Infection cases: 'suspected,' 'probable,' and 'confirmed'
- Approximately weekly infection cases: data recorded and reported have not been completely regular
- Obvious data inaccuracies such as negative weekly numbers of infections were corrected prior to data analysis.

Irregularities in Data

Odd Instance at Sensible Time Point:

- Instance emerged around time when intervention showed an effect in week of Nov 4, 2014 (week 33)
- Weekly infection numbers peaked in week of Nov 4, 2014
- Time of intervention in analysis obviously is in week 33
- In week 33, total (weekly) number of infections: 13268 (4052)
- In week 34, total number of infections: < 13268.
- More plausible weekly infection numbers were imputed for data analysis in weeks 33 and 34

Level Shift Into Subcritical Regime

Results Shown

- Phase transition between supercritical and subcritical behaviors is delineated by $\mathbf{E}(V_{\kappa_t}) = \mathbf{E}[\log(\kappa_t)] = 0$

- **'Escape from Supercritical Phase for Level Shift'**

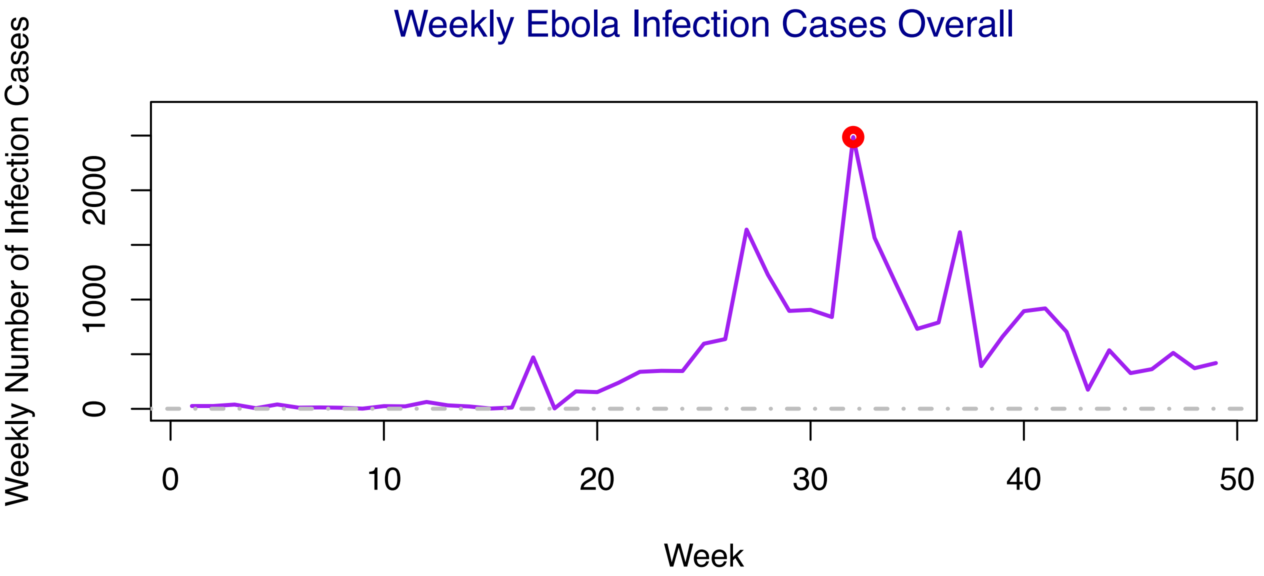
  Assume that process prior to intervention is stationary and supercritical.

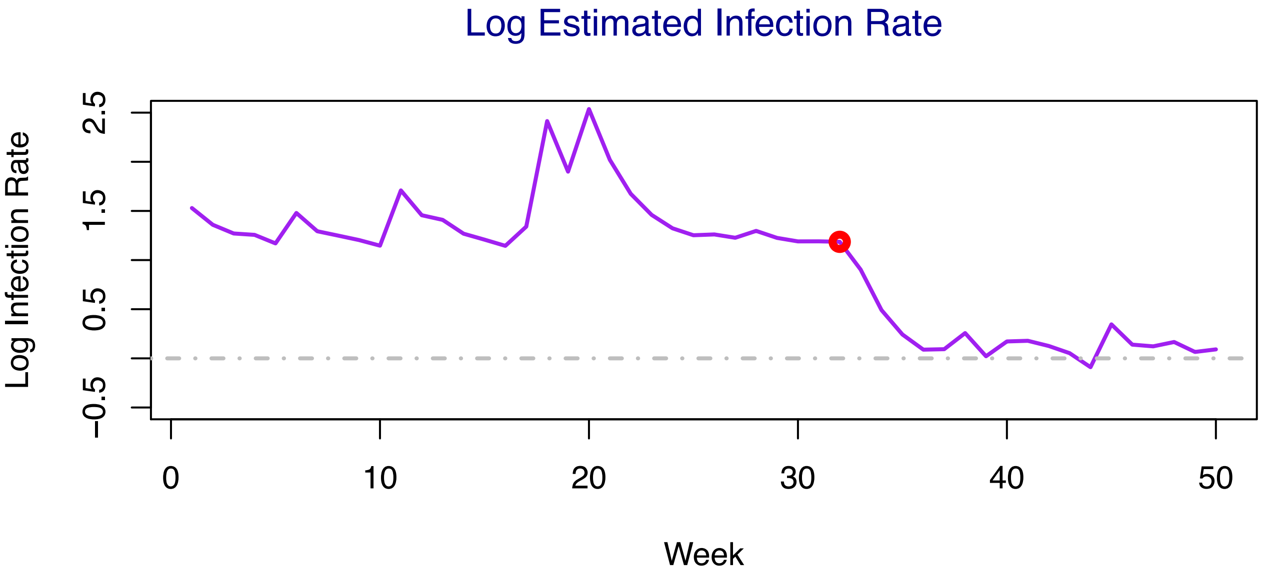
  If effect size $\nu < 0$ is such that $\mathbf{E}(\kappa_{\tau+h_*}) < 1$ for some $h_* \geq 0$, then process after time $\tau + h_*$ is subcritical.

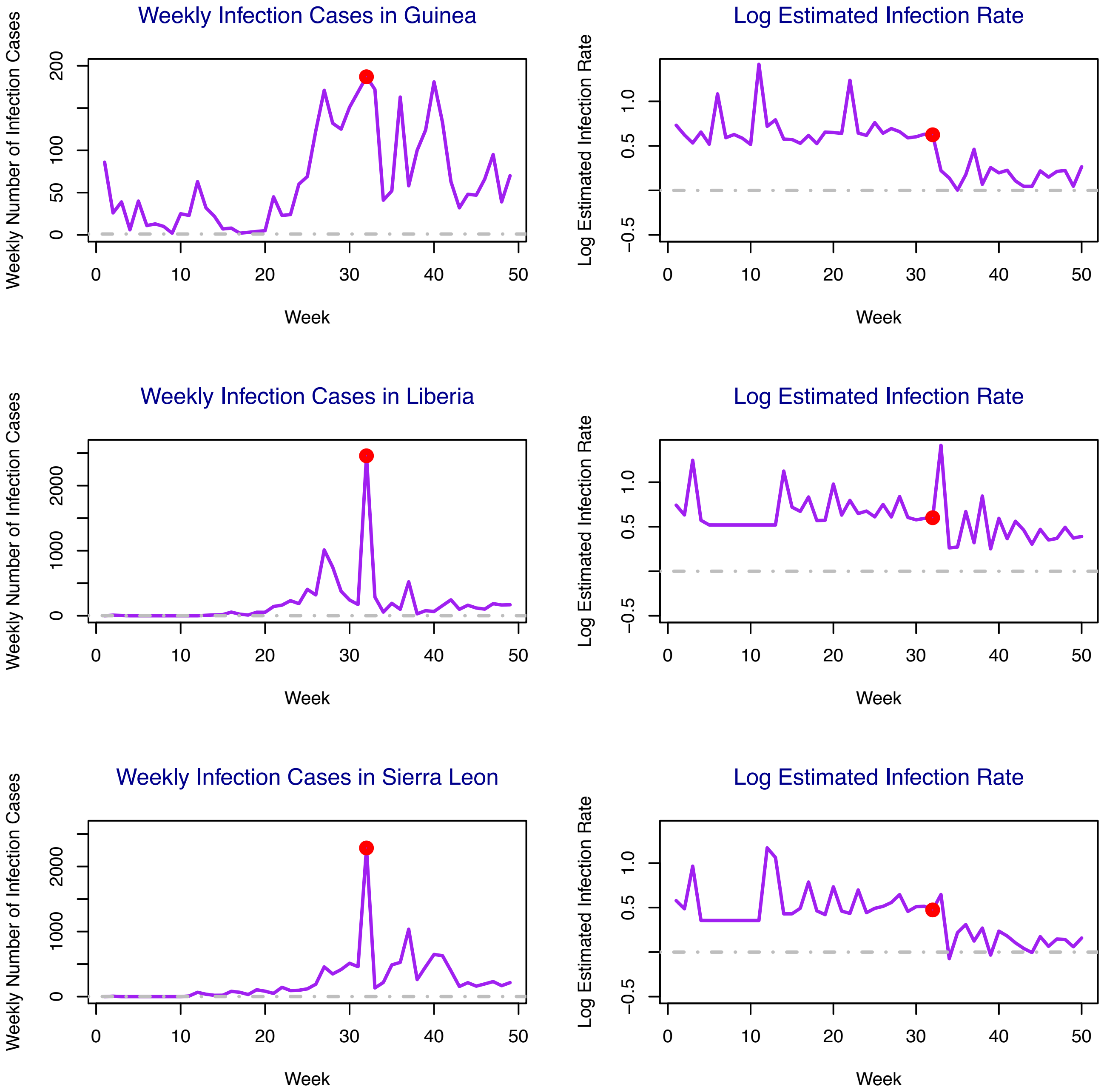
Intervention Analysis: Results

Model Fit

- Fit a Poisson GARCE(1, 1) BP with level shift intervention at $\tau = 33$
- $n = 49$ observations for combined regions Guinea, Liberia, and Sierra Leon
- MLEs: $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\nu}) = (1.59, 0.20, 0.46, -1.21)$
- Estimated mean level prior to intervention: $\hat{\mu} = 4.61$
- Estimated mean level after intervention: $\hat{\tilde{\mu}}_1 = 1.12,$
- Estimate for $\mathbf{E}(V_{\kappa_t}) = \mathbf{E}[\log(\kappa_t)]$: $\hat{v}_\kappa = 0.193.$


Weekly Ebola Infection Cases Overall
Weekly Number of Infection Cases
0
1000
2000
0
10
20
30
40
50
Week


Log Estimated Infection Rate
Log Infection Rate
−0.5
0.5
1.5
2.5
0
10
20
30
40
50
Week


Weekly Infection Cases in Guinea
Weekly Number of Infection Cases
0
50
100
200
0
10
20
30
40
50
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
−0.5
0.5
1.0
0
10
20
30
40
50
Week
Weekly Infection Cases in Liberia
Weekly Number of Infection Cases
0
1000
2000
0
10
20
30
40
50
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
−0.5
0.5
1.0
0
10
20
30
40
50
Week
Weekly Infection Cases in Sierra Leon
Weekly Number of Infection Cases
0
1000
2000
0
10
20
30
40
50
Week
Log Estimated Infection Rate
Log Estimated Infection Rate
−0.5
0.5
1.0
0
10
20
30
40
50
Week

Intervention Analysis: Conclusions

Conclusions

- Weekly infection rate dropped sharply after intervention from an average level of $3.5 - 4$ to level of ≈ 1. One instance was clearly below 1.
- By end of February of 2015, infection rate has leveled off and oscillates around number above 1.
- Thus, overall infection cases moved in supercritical phase but near phase transition to subcritical phase.

Estimates of Extinction Probability and Time to Extinction

- We saw that non-certain extinction is only possible for a supercritical process.

  The extinction probability q can be obtained by numerically solving an equation that depends on model parameters.

- For subcritical and critical processes, extinction is certain almost surely relative to environment (aside from degen. case for critical process).

  Time to extinction can be simulated employing the BP model with intervention effect.

THANK YOU

THANK YOU !!!

List of References

References

- K.B. Athreya, S. Karlin, On branching processes with random environments. I. Extinction probabilities, Ann. Math. Statist. 42 (1971) 1499–1520.
- K.B. Athreya, S. Karlin, Branching processes with random environments. II. Limit theorems, Ann. Math. Statist. 42 (1971) 1843–1858.
- K.B. Athreya, P. Ney, Branching Processes. Reprint of the 1972 original (Springer, New York), Dover Publications, Mineola, NY, 2004.
- T.E. Harris, The Theory of Branching Processes. Die Grundlehren der Mathematischen Wissenschaften, 119, Springer-Verlag, Berlin, 1963.

- I. Hueter, Branching processes in generalized autoregressive conditional environments, Preprint (2014).
- I. Hueter, Interventions in GARCE branching processes, Preprint (2015).
- W.L. Smith, W.E. Wilkinson, On branching processes in random environments, Ann. Math. Statist. 40 (1969) 814–827.
- W.L. Smith, W.E. Wilkinson, Branching processes in Markovian environments, Duke Math. J. 38 (1971) 749–763.