Arizona State University

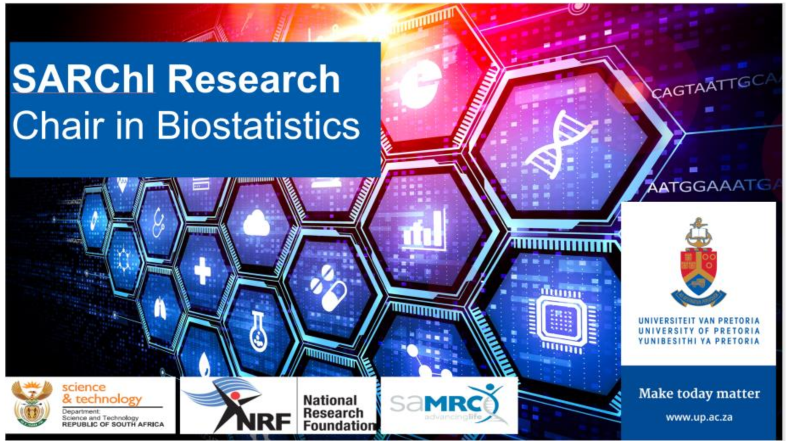
Susceptible-Infected-Recovered Models: Malaria in Ethiopia

International Meetings on Differential Equations & Their Applications.
Institute of Mathematics, Lodz University of Technology, Łódź, Poland

Din Chen, Ph.D.
Elected Member | Academy of Science of South Africa
Elected Fellow | American Statistical Association

SARCHI Research Chair & Extraordinary Professor in Biostatistics
Department of Statistics, University of Pretoria | South Africa

Executive Director & Professor in Biostatistics | College of Health Solutions
Senior Global Futures Scientist | Julie Ann Wrigley Global Futures Laboratory
Arizona State University, Phoenix, USA



Outline

1. Mathematical SIR Model
2. SIR to SEIRD Model
3. Statistical SEIRD Model
4. Take Home Messages

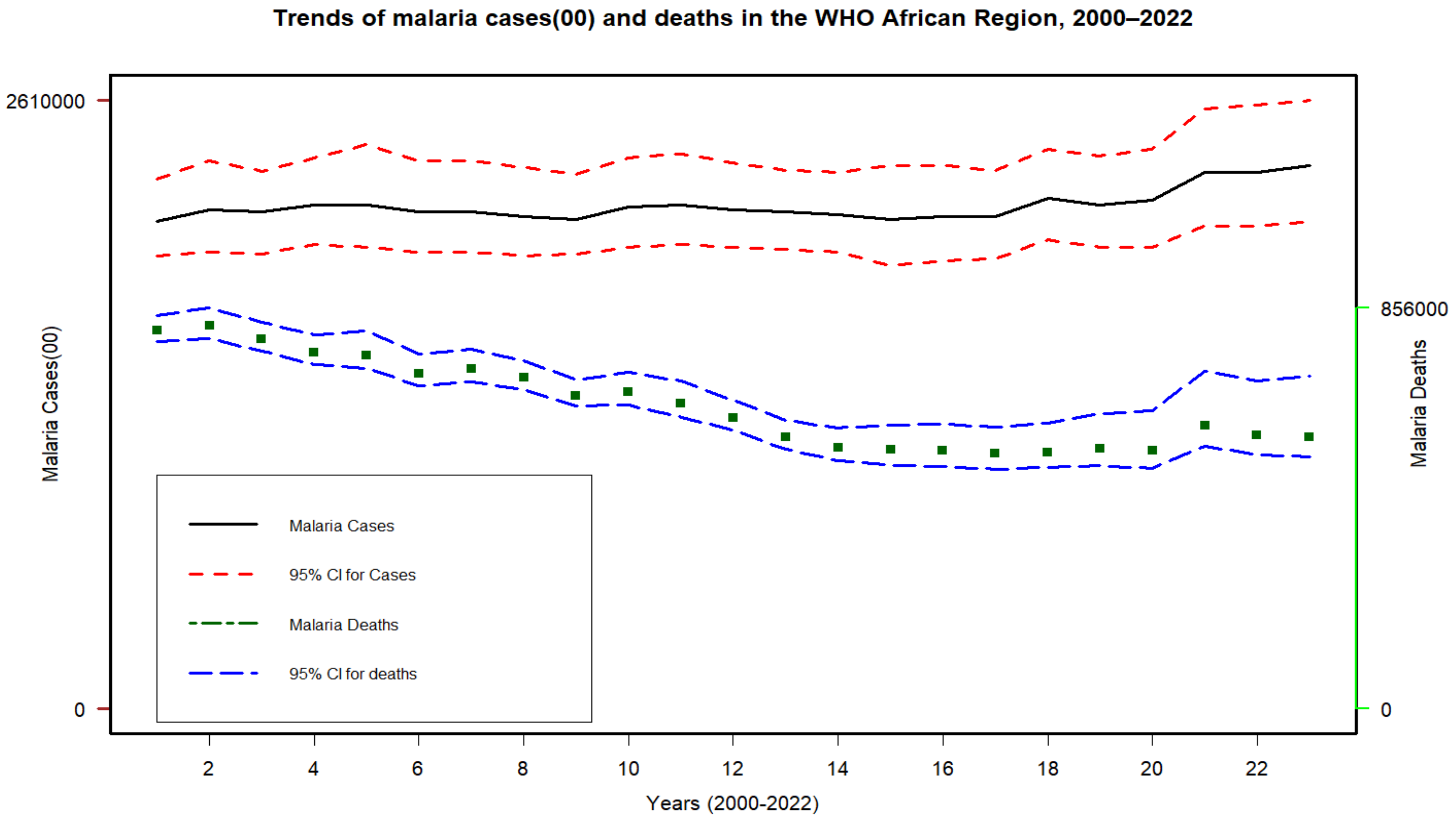
Outline

1. Mathematical SIR Model
2. SIR to SEIRD Model
3. Statistical SEIRD Model
4. Take Home Messages

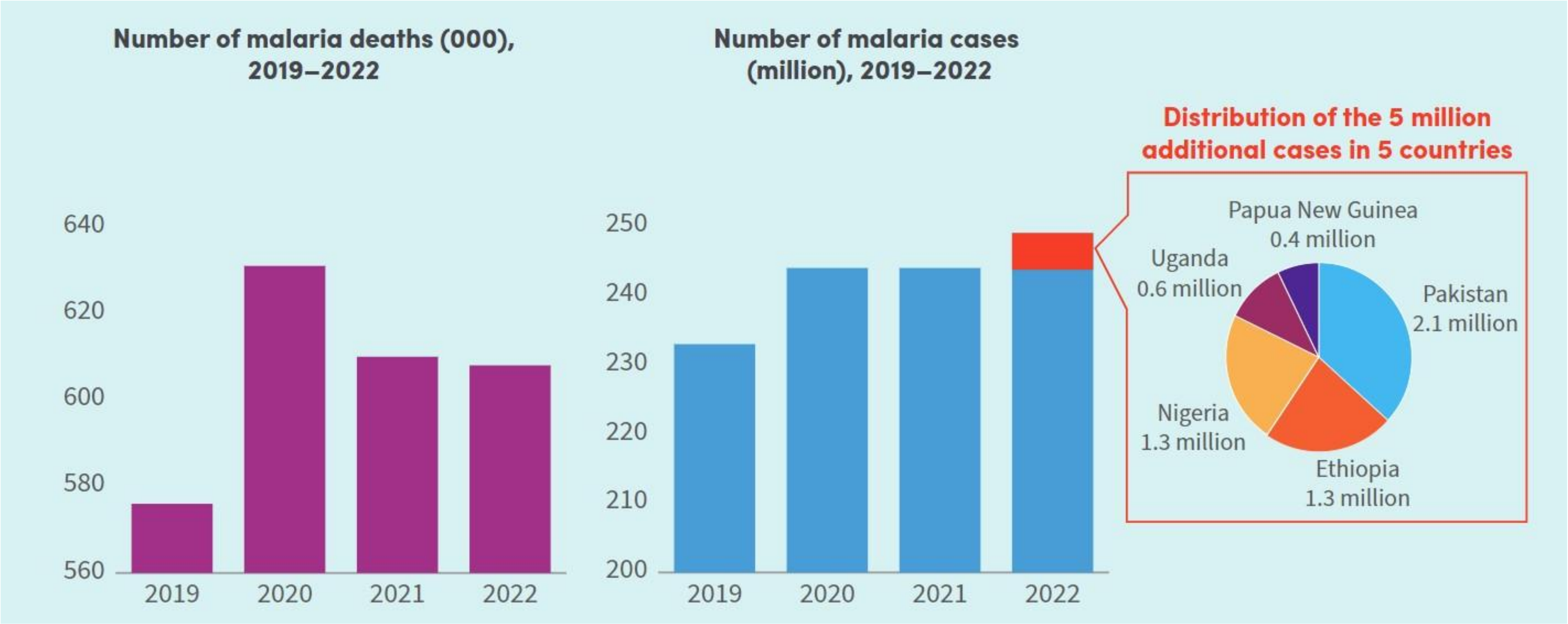
What is a Mathematical Model?

- A mathematical description for the real-world
- Only for some specific quantitative features of the real-world, ignores others for simplification
- With purpose to improve understanding of the real-world

Estimated Malaria Cases & Deaths 2000 to 2022 (WHO 2023)



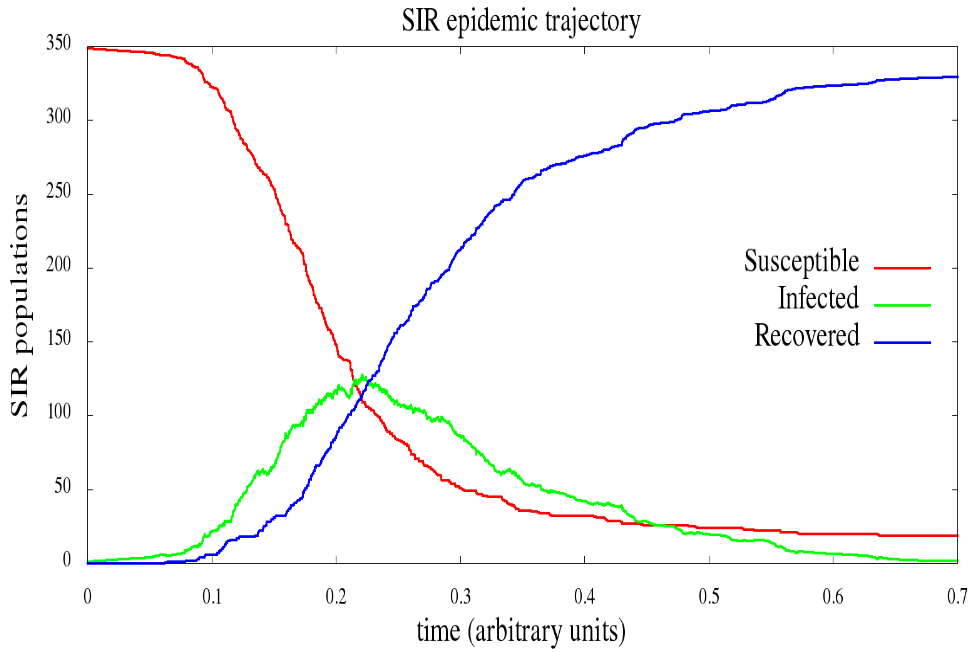
Estimated Malaria Cases & Deaths 2000 to 2022 (WHO 2023)



The SIR Epidemic Model

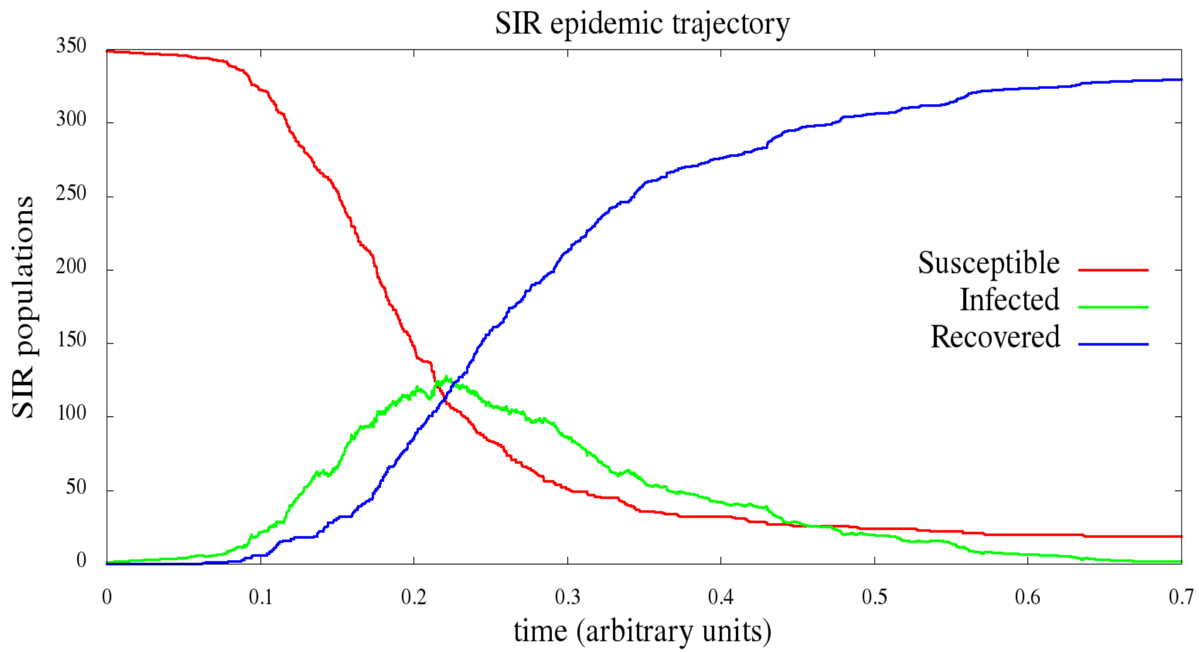
- Epidemiological model to model the dynamics of infectious diseases
- First studied by Kermack & McKendrick, 1927 (https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118)
- Consider a disease spread by contact with infected individuals.
- Individuals recover from the disease & gain further immunity from it.
- S = fraction of *susceptibles* in a population
- I = fraction of *infecteds* in a population
- R = fraction of *recovereds* in a population

$$S + I + R = N$$



The SIR Epidemic Model

- The SIR model is a compartmental model to categorize populations into compartments
- The SIR model then model all compartments with differential models and their interactions with the time trajectories & disease progressions.



The SIR Model in Mathematics Equations

$$\frac{dS}{dt} = -\beta \cdot \frac{S \cdot I}{N}$$

$$\frac{dI}{dt} = \beta \cdot \frac{S \cdot I}{N} - \gamma \cdot I$$

$$\frac{dR}{dt} = \gamma \cdot I$$

β: **Transmission rate** (contact rate × probability of transmission per contact).

γ: **Recovery rate** (inverse of average infectious period).

$R_0 = \frac{\beta}{\gamma}$: Basic reproduction number — the average number of new infections caused by one infected individual in a fully susceptible population.

Outline

1. Mathematical SIR Model
2. SIR to SEIRD Model
3. Statistical SEIRD Model
4. Take Home Messages

Enhancing the SIR Model

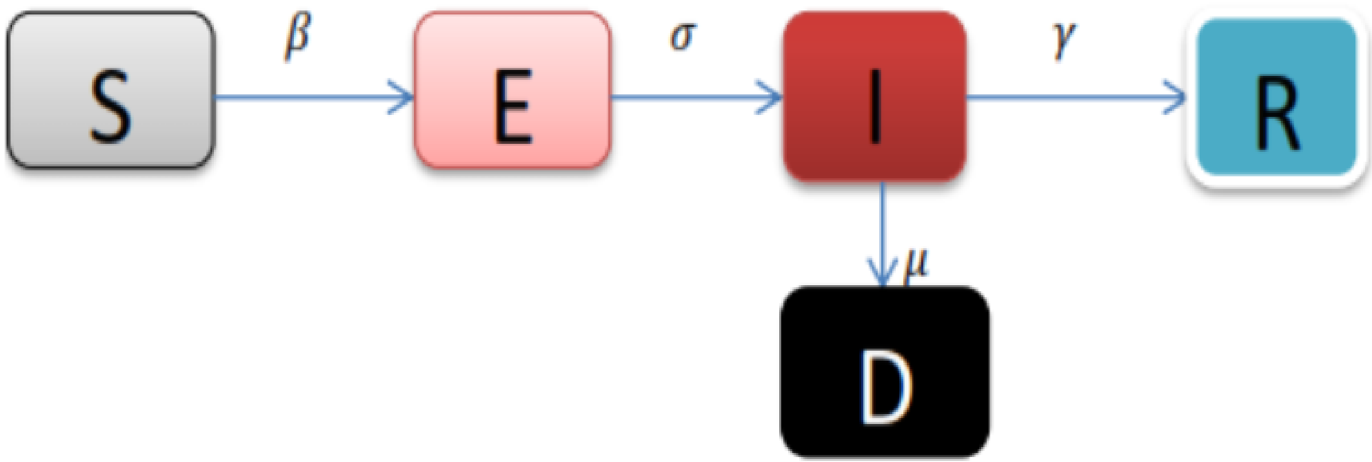
- **SEIR**: Consider **E***xposed* (Infected but not yet infectious; incubation period) into SIR.

- **SEIRD**: Consider **D**eath into SEIR.

- Can consider age differences, multiple types of transmission, geographic spread, etc.

The SEIRD Model in Mathematics Equations

$$\frac{dS}{dt} = -\beta \cdot \frac{S \cdot I}{N}$$
$$\frac{dE}{dt} = \beta \cdot \frac{S \cdot I}{N} - \sigma \cdot E$$
$$\frac{dI}{dt} = \sigma \cdot E - \gamma \cdot I - \mu \cdot I$$
$$\frac{dR}{dt} = \gamma \cdot I$$
$$\frac{dD}{dt} = \mu \cdot I$$



β: Transmission rate

σ: Incubation rate (inverse of incubation period)

γ: Recovery rate

μ: Mortality rate

$N = S + E + I + R + D$: Total population

Outline

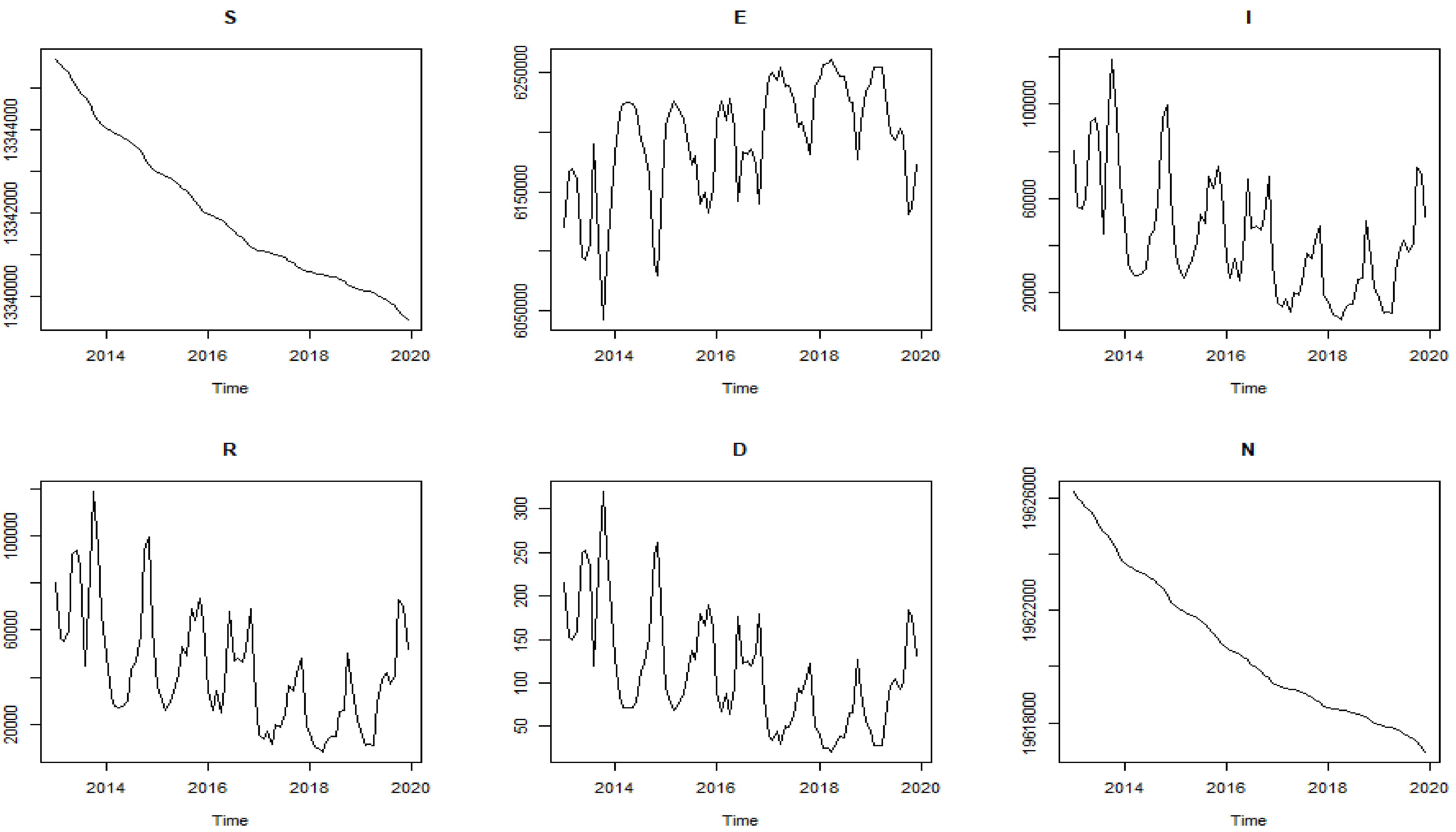
1. Mathematical SIR Model
2. SIR to SEIRD Model
3. Statistical SEIRD Model
4. Take Home Messages

Data Available

Ethiopia Data:

- ➢Time Series for S/E/I/R/D/N
- ➢Time t = 1 to T (84 months)
- ➢From 2013 to 2019 (7 years)

	A	B	C	D	E	F	G	H
1	Year	Month	I	D	R	N	S	E
2	2013	January	80222	215	80007	19626000	13345680	6120091
3	2013	February	56627	152	56475	19625848	13345577	6167169
4	2013	March	55458	149	55309	19625699	13345475	6169457
5	2013	April	59688	160	59528	19625539	13345366	6160957
6	2013	May	92702	249	92453	19625290	13345197	6094938
7	2013	June	93915	252	93663	19625037	13345025	6092434
8	2013	July	87110	234	86876	19624803	13344866	6105951
9	2013	August	44757	120	44637	19624683	13344785	6190505
10	2013	September	85334	229	85105	19624454	13344629	6109386
11	2013	October	118797	319	118478	19624135	13344412	6042448
12	2013	November	94620	254	94366	19623881	13344239	6090656
13	2013	December	67232	181	67051	19623700	13344116	6145301
14	2014	January	49083	128	48955	19623572	13344029	6181505
15	2014	February	32340	85	32255	19623487	13343971	6214920
16	2014	March	27688	72	27616	19623414	13343922	6224189
17	2014	April	27400	72	27328	19623343	13343873	6224741
18	2014	May	27566	72	27494	19623271	13343824	6224387
19	2014	June	29762	78	29684	19623193	13343771	6219976
20	2014	July	43226	113	43113	19623080	13343694	6193047
21	2014	August	47076	123	46953	19622956	13343610	6185317
22	2014	September	57978	152	57826	19622804	13343507	6163493
23	2014	October	94259	247	94012	19622558	13343339	6090947
24	2014	November	99835	261	99574	19622296	13343162	6079726
25	2014	December	60951	160	60791	19622137	13343053	6157341



S
E
I
R
D
N
Time
2014
2016
2018
2020
13340000
13342000
13344000
6050000
6150000
6250000
20000
60000
100000
50
100
150
200
250
300
19618000
19622000
19626000

Data Summary

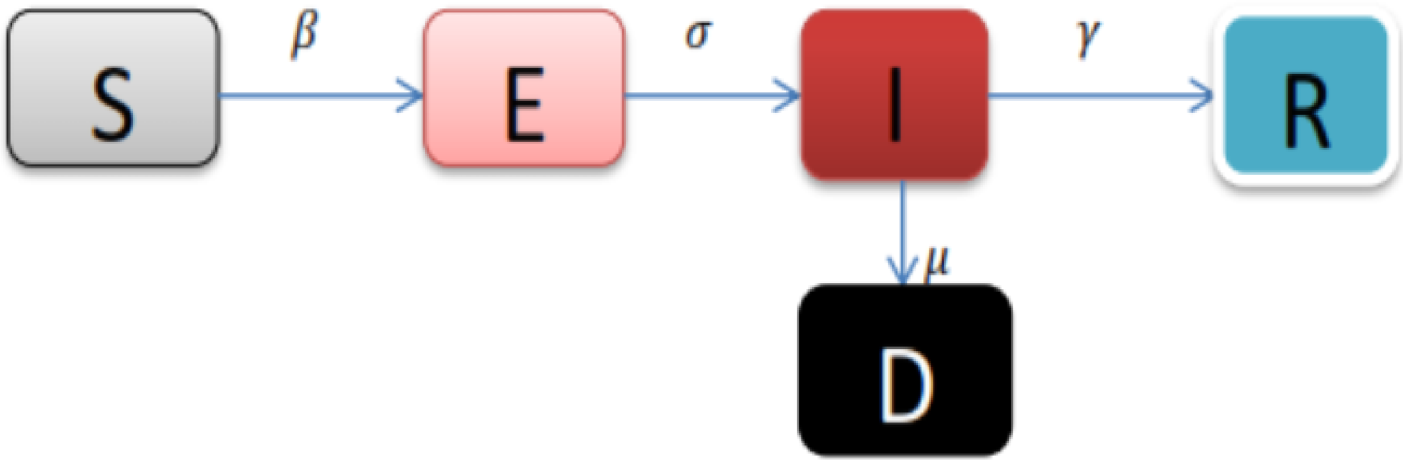
➢Data Summary:

	S	E	I	R	D	N
Mean	13,341,958	6,193,572	42,553	42,443	111	19,620,638
SD	1,689	49,060	24,770	24,704	66	2,521

	ΔS	ΔE	ΔI	ΔR	ΔD
Mean	-75.12	637.16	-336.76	-335.75	-1.01
SD	44.93	32634.59	16335.45	16292.64	42.82

The SEIRD Model in Mathematics Equations (ODE)

$$\frac{dS}{dt} = -\beta \cdot \frac{S \cdot I}{N}$$
$$\frac{dE}{dt} = \beta \cdot \frac{S \cdot I}{N} - \sigma \cdot E$$
$$\frac{dI}{dt} = \sigma \cdot E - \gamma \cdot I - \mu \cdot I$$
$$\frac{dR}{dt} = \gamma \cdot I$$
$$\frac{dD}{dt} = \mu \cdot I$$



$$N = S + E + I + R + D.$$

Statistical Considerations of
The Stochastic SEIRD Model (SDE)

$$
\begin{aligned}
dS_t &= -\frac{\beta S_t I_t}{N_t} dt + \sigma_S dW_{S_t} \\
dE_t &= \left[\frac{\beta S_t I_t}{N_t} - \sigma E_t\right] dt + \sigma_E dW_{E_t} \\
dI_t &= [\sigma E_t - (\gamma + \mu) I_t]\, dt + \sigma_I dW_{I_t} \\
dR_t &= \gamma I_t dt + \sigma_R dW_{R_t} \\
dD_t &= \mu I_t dt + \sigma_D dW_{D_t}
\end{aligned}
$$

- Drift terms (from the ODE): deterministic changes.
- Diffusion terms: $\sigma_X\, dW_X$ where σ_X controls stochastic intensity and dW_X is a Wiener process (Brownian motion).
- Each dW_X can be independent or correlated (e.g., shared noise for epidemic shocks).

SDE System to Difference Equation Systems:
Still the SEIRD Model (Euler-Maruyama Method)

For Statistical Estimation and Inference

$$S_{t+\Delta t} = S_t - \beta \frac{S_t I_t}{N} \Delta t + \sigma_S \cdot \Delta W_S$$

$$E_{t+\Delta t} = E_t + \left(\beta \frac{S_t I_t}{N} - \alpha E_t \right) \Delta t + \sigma_E \cdot \Delta W_E$$

$$I_{t+\Delta t} = I_t + (\alpha E_t - \gamma I_t - \mu I_t) \Delta t + \sigma_I \cdot \Delta W_I$$

$$R_{t+\Delta t} = R_t + \gamma I_t \Delta t + \sigma_R \cdot \Delta W_R$$

$$D_{t+\Delta t} = D_t + \mu I_t \Delta t + \sigma_D \cdot \Delta W_D$$

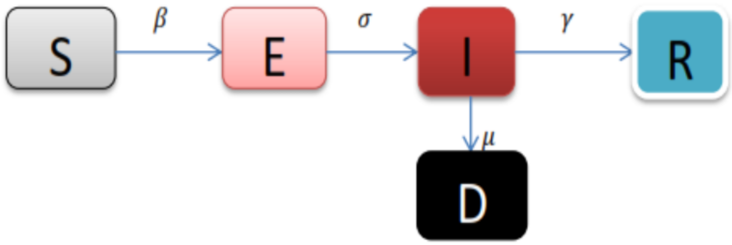
- σ_X: magnitude of noise for compartment X
- $\Delta W_X \sim \mathcal{N}(0, \Delta t)$: random noise (Wiener increment)

Weighted Least-Squared Estimation:

Still the SEIRD Model

$$L(\beta,\sigma,\gamma,\mu,\sigma_S,\sigma_E,\sigma_I,\sigma_R,\sigma_D) = \sum_{t=1}^{T-1}\frac{\left(\Delta S_t + \frac{\beta S_t I_t}{N_t}\right)^2}{\sigma_S^2} + \sum_{t=1}^{T-1}\frac{\left(\Delta E_t - \frac{\beta S_t I_t}{N_t} + \sigma E_t\right)^2}{\sigma_E^2} + \sum_{t=1}^{T-1}\frac{(\Delta I_t - \sigma E_t - (\gamma+\mu)I_t)^2}{\sigma_I^2} + \sum_{t=1}^{T-1}\frac{(\Delta R_t - \gamma I_t)^2}{\sigma_R^2} + \sum_{t=1}^{T-1}\frac{(\Delta D_t - \mu I_t)^2}{\sigma_D^2}$$

Loss Function to be minimized for parameter estimation:



The Maximum Likelihood Estimation:

The SEIRD Model

$$L(\beta, \sigma, \gamma, \mu, \sigma_S, \sigma_E, \sigma_I, \sigma_R, \sigma_D) = \prod_{t=1}^{T-1} dnorm\left(\Delta S_t, -\frac{\beta S_t I_t}{N_t}, \sigma_S\right) \times$$

$$\prod_{t=1}^{T-1} dnorm\left(\Delta E_t, \frac{\beta S_t I_t}{N_t} - \sigma E_t, \sigma_E\right) \times$$

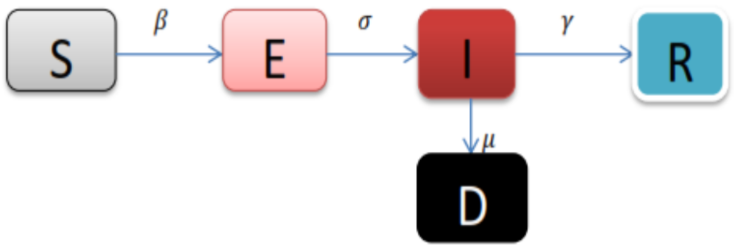
$$\prod_{t=1}^{T-1} dnorm\left(\Delta I_t, \sigma E_t - (\gamma + \mu) I_t, \sigma_I\right) \times$$

$$\prod_{t=1}^{T-1} dnorm\left(\Delta R_t, \gamma I_t, \sigma_R\right) \times$$

$$\prod_{t=1}^{T-1} dnorm\left(\Delta D_t, \mu I_t, \sigma_I\right)$$

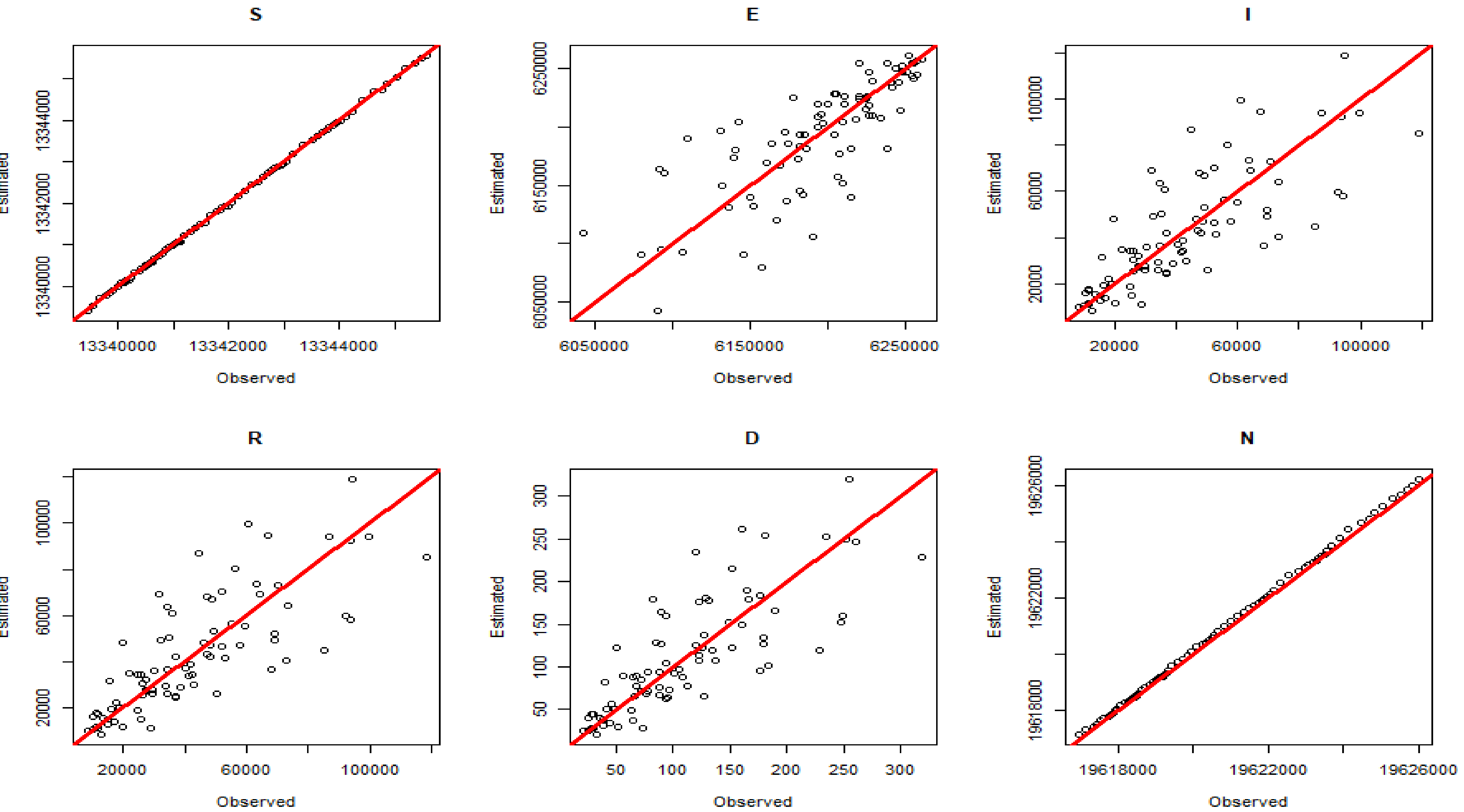
Likelihood Function to be maximized for parameter estimation:

$$dnorm(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

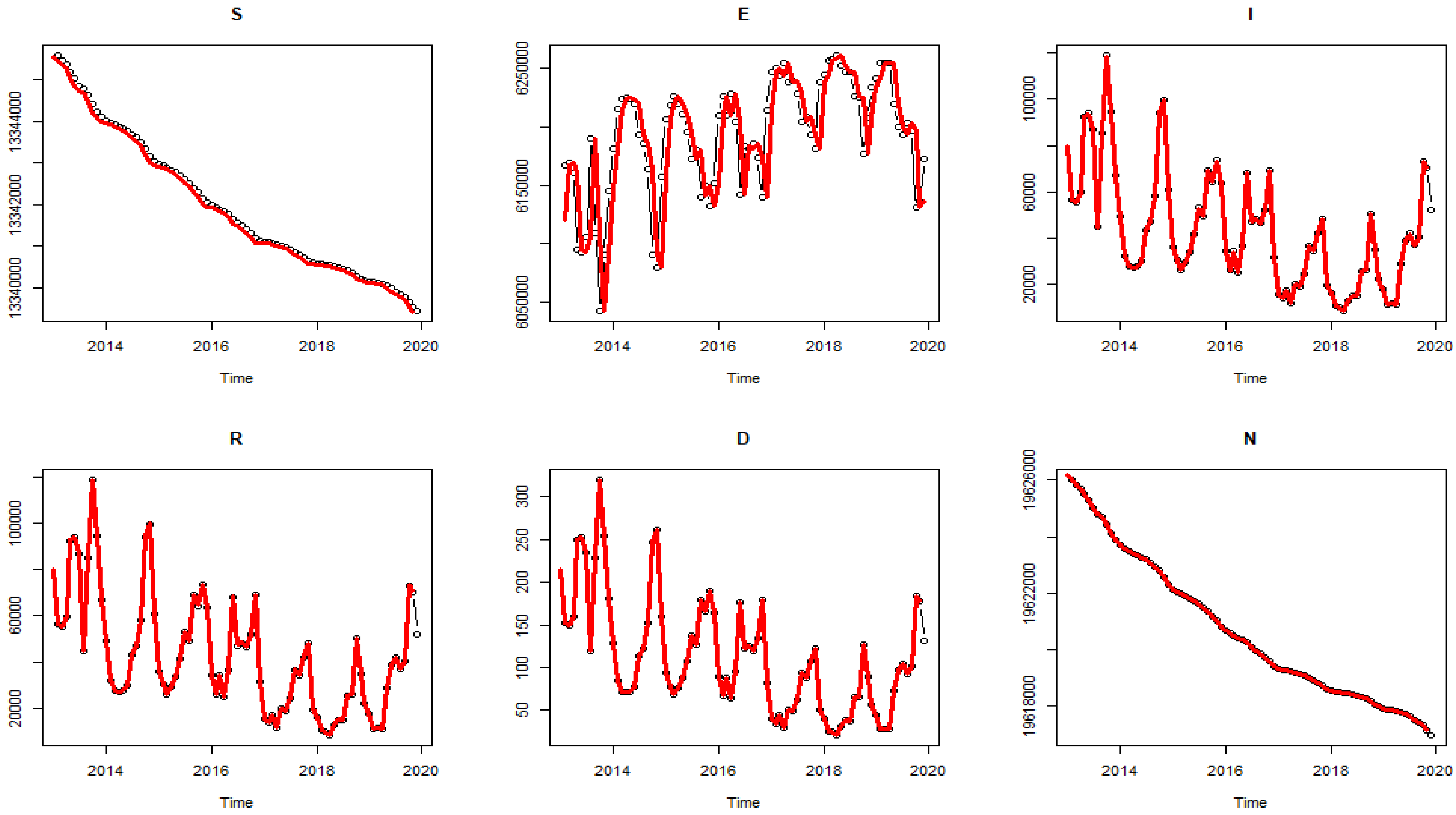


Model Fitting Diagnostics

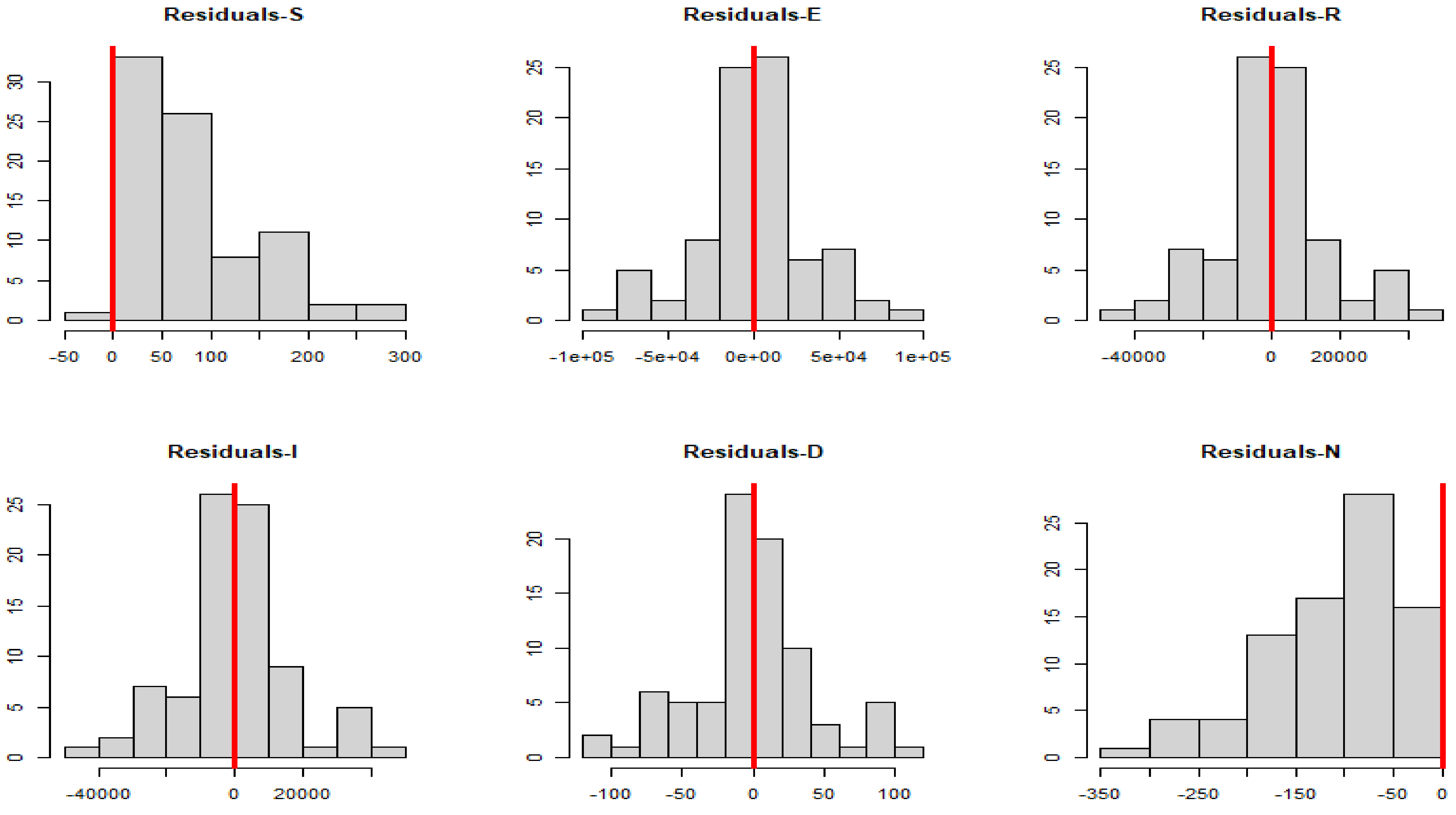
- ✓ **Observed vs Fitted Scatterplot**
  - 1-1 Alignment!
- ✓ **Observed vs.. Fitted Time Series**
  - Time Series Alignment
- ✓ **Residual Plots**
  - Normally Distributed Residuals, White Noises!



S
E
I
R
D
N
Observed
Estimated
13340000
13342000
13344000
6050000
6150000
6250000
20000
60000
100000
50
100
150
200
250
300
19618000
19622000
19626000



S
13340000
13342000
13344000
2014
2016
2018
2020
Time
E
6050000
6150000
6250000
2014
2016
2018
2020
Time
I
20000
60000
100000
2014
2016
2018
2020
Time
R
20000
60000
100000
2014
2016
2018
2020
Time
D
50
100
150
200
250
300
2014
2016
2018
2020
Time
N
19618000
19622000
19626000
2014
2016
2018
2020
Time



Residuals-S
0
5
10
15
20
25
30
-50
0
50
100
200
300
Residuals-E
0
5
10
15
20
25
-1e+05
-5e+04
0e+00
5e+04
1e+05
Residuals-R
0
5
10
15
20
25
-40000
0
20000
Residuals-I
0
5
10
15
20
25
-40000
0
20000
Residuals-D
0
5
10
15
20
-100
-50
0
50
100
Residuals-N
0
5
10
15
20
25
-350
-250
-150
-50
0

Statistical Estimation & Inference

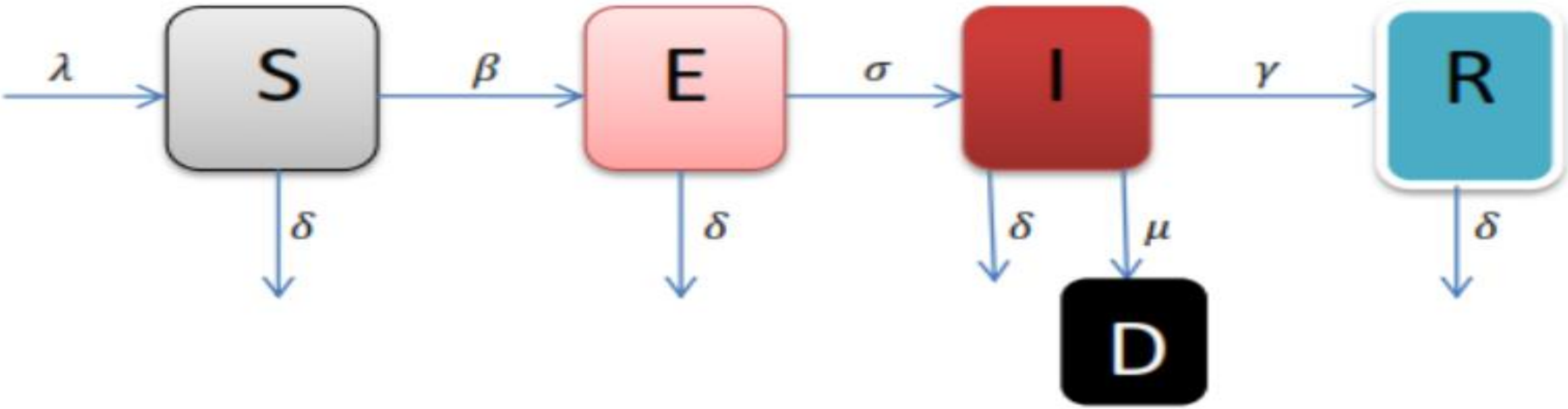
Parameter	Estimate	
β	2.474 ($\times 10^{-3}$)	Transmission Rate
σ	2.113 ($\times 10^{-56}$)	Incubation Rate
γ	3.060 ($\times 10^{-3}$)	Recovery Rate
μ	7.667 ($\times 10^{-15}$)	Mortality Rate
σ_S	183.233	Standard Deviation (Stochastic Intensity) of S
σ_E	32617.057	Standard Deviation (Stochastic Intensity) of E
σ_I	16480.639	Standard Deviation (Stochastic Intensity) of I
σ_R	16304.223	Standard Deviation (Stochastic Intensity) of R
σ_D	247.749	Standard Deviation (Stochastic Intensity) of D
$R_0 = \frac{\beta}{\gamma + \mu}$	0.808	Basic Reproduction Number

Notes (Each infected person, on average, transmits the disease to less than one other person)

- **Expected behavior:** The number of infection will decline over time, and the disease will die out without an epidemic
- **No exponential growth:** Instead of spreading, the infection may fades out.

Other SEIRD Models (1)

- Assuming dynamic demography with existence of **birth rate λ** which flows to the susceptible populations & **natural death rate δ** which simultaneously deducted from each compartment, causing each proportion of the population to vary over time



Other SEIRD Models (1)

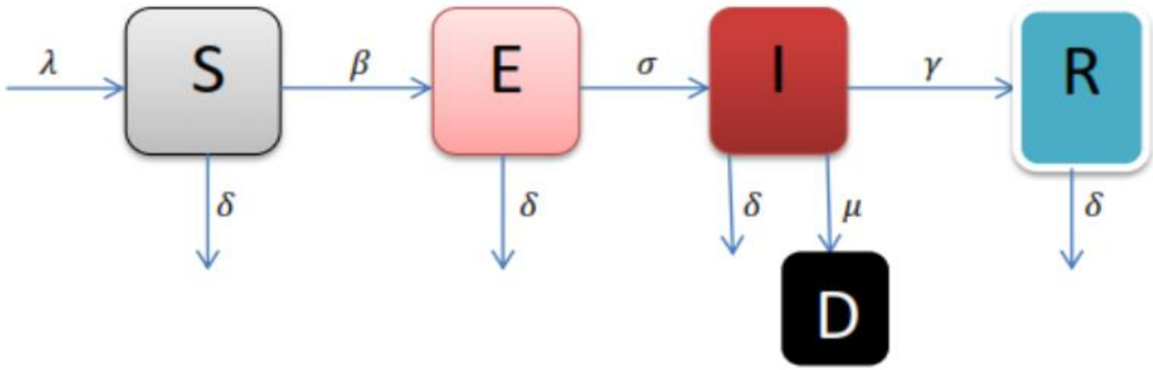
- $$\frac{dS}{dt} = \lambda N - \frac{\beta SI}{N} - \delta S$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - (\sigma + \delta)E$$

$$\frac{dI}{dt} = \sigma E - (\gamma + \mu + \delta)I$$

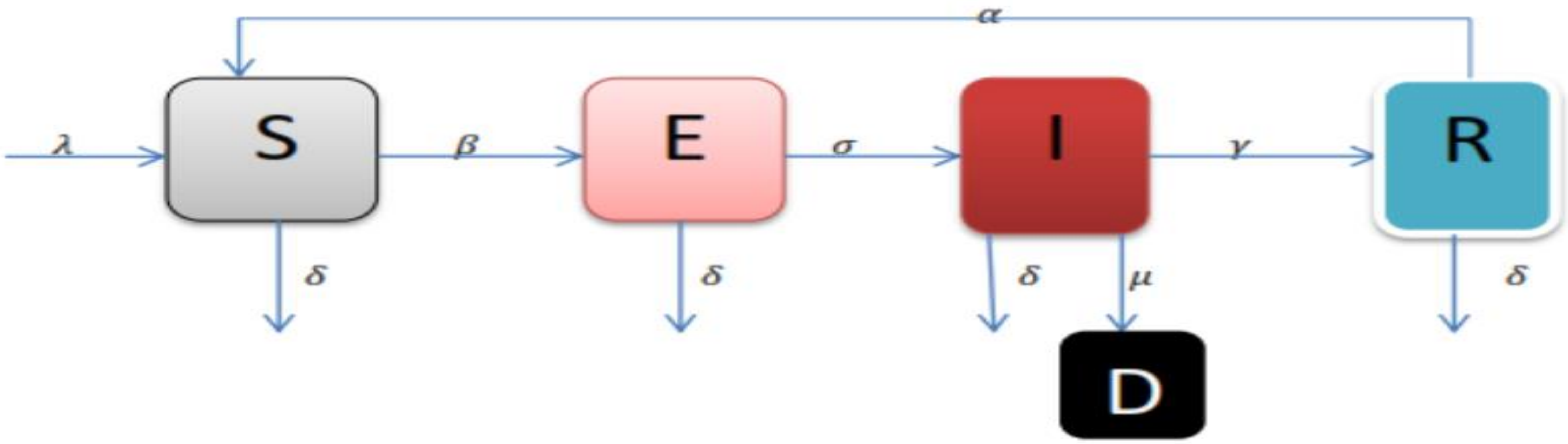
$$\frac{dR}{dt} = \gamma I - \delta R$$

$$\frac{dD}{dt} = \mu I$$



Other SEIRD Models (2)

- The recovered population will return back to the susceptible populations:



Other SEIRD Models (2)

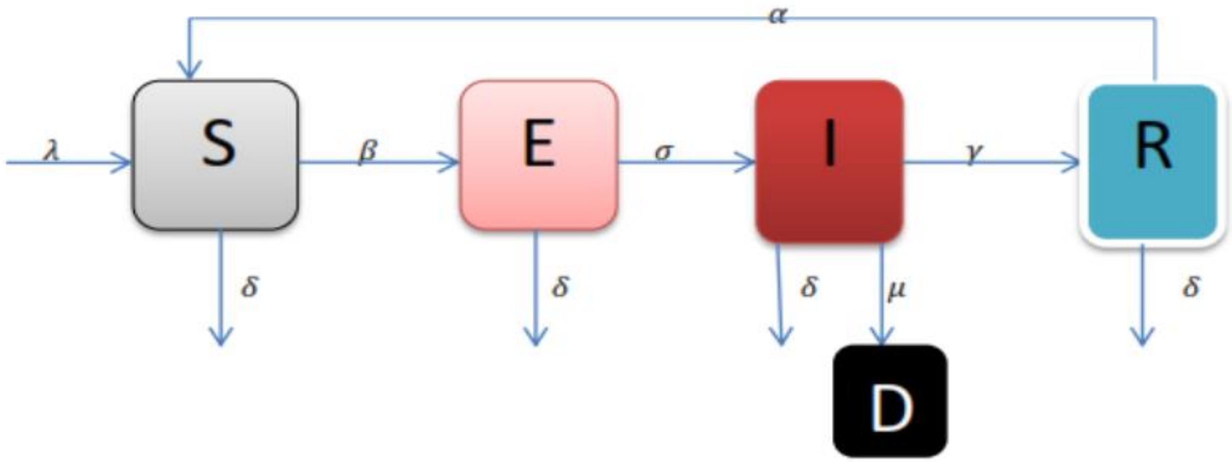
$$\frac{dS}{dt} = \lambda N - \frac{\beta SI}{N} + \alpha R - \delta S$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - (\sigma + \delta)E$$

$$\frac{dI}{dt} = \sigma E - (\gamma + \mu + \delta)I$$

$$\frac{dR}{dt} = \gamma I - \delta R$$

$$\frac{dD}{dt} = \mu I$$



Outline

1. Mathematical SIR Model
2. SIR to SEIRD Model
3. Statistical SEIRD Model
4. Take Home Messages

Take Home Messages

- ✓ **SIR and its Extension for Disease Modeling**
  - Many Mathematical Models Available!
- ✓ **Statistical Considerations**
  - Real-World Data Compilation (Time Consuming)!
  - Parameter Estimation with Weighted Least Squared Estimation
  - Parameter Estimation with Maximum Likelihood Estimation
- ✓ **Prediction with/without Interventions**
  - Simulate Intervention Strategies after Estimation!

Future Research

✓More SIR Models (Application-Based)

- Extension of Existing SIR-related Models!
- Spatial-Temporal SIR Models (such as, COVID-19)
- Open Population SIR Model

✓Statistical Considerations

- Parameter Estimation and Statistical Inference
- Approximate Bayesian Computation (ABC):
  - Likelihood-free method—ideal for the complex SDE models.
- Software Development (such as open-source package in R)

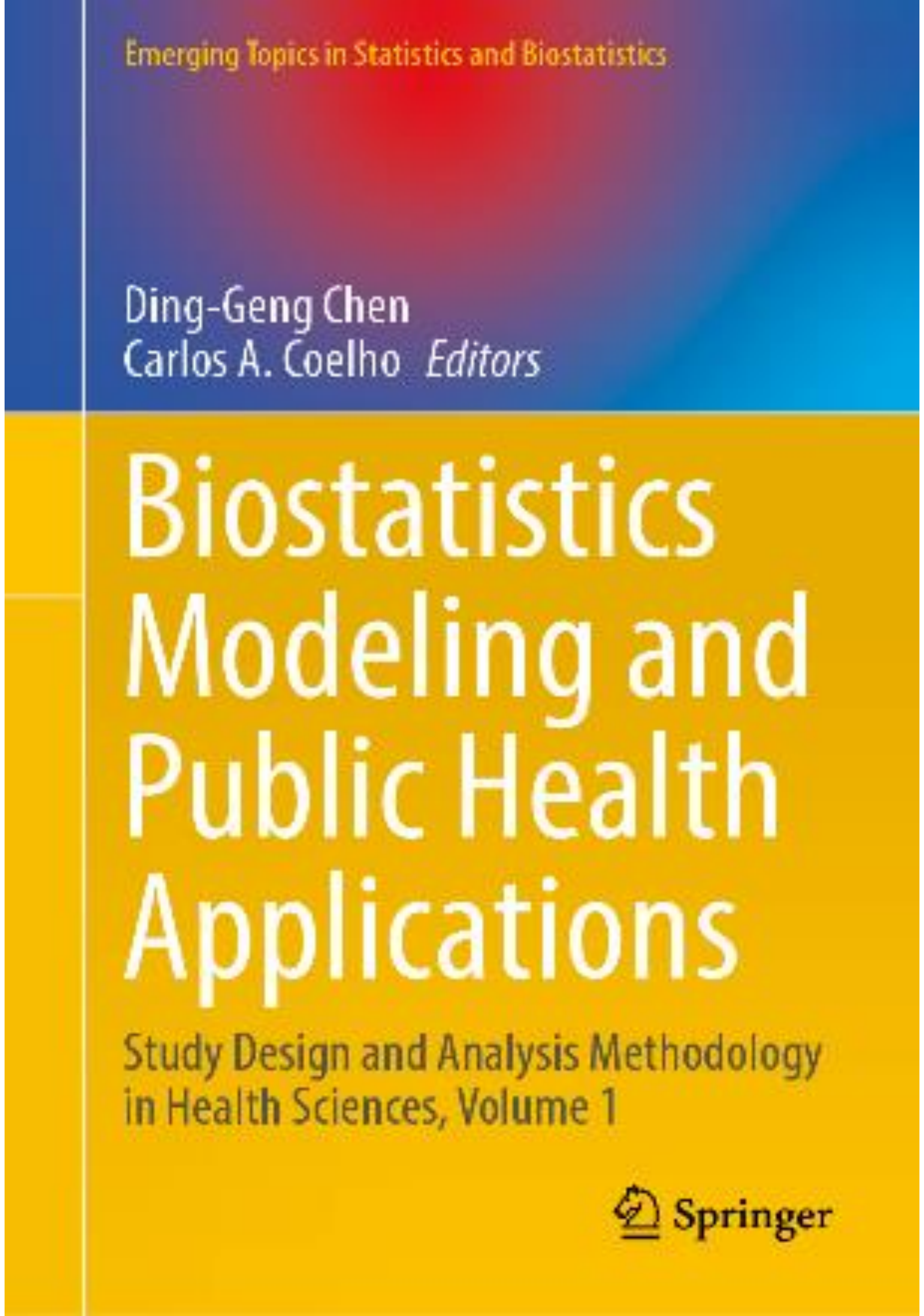
✓Predictions with/without Interventions

- Simulate Intervention Strategies after Model Estimation!

<u>Chapter 8:</u>

SEIRD Mathematical Modeling of Malaria Transmission Dynamics in Ethiopia, 193-219, by Belay, D. B. **Chen, D. G.** and Matintu, S. A.

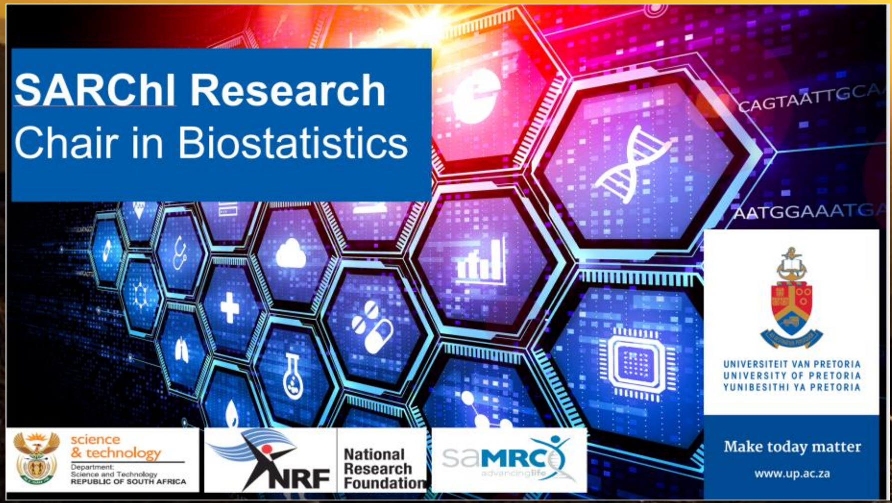
SOURCE: Chen DG, Coelho CA. Biostatistics Modeling and Public Health Applications. Springer Nature; 2024.



More References

- J. D. Murray, *Mathematical Biology*, Springer-Verlag, 1989.
- Allyn Jackson, Modeling the Aids Epidemic, *Notices of the American Mathematical Society*, 36:981-983, 1989.
- O. Diekmann & A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases*, Wiley, 2000.
- Matt Keeling, *The Mathematics of Diseases*, http://plus.maths.org, 2004.
- Ottar N. Bjornstad, Epidemics: Models and Data Using R. 2014. Springer.

Thank You!



SARChI Research
Chair in Biostatistics
CAGTAATTGCA
AATGGAAATGA
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Make today matter
www.up.ac.za
science & technology
Department: Science and Technology
REPUBLIC OF SOUTH AFRICA
NRF
National Research Foundation
saMRC
advancing life