The Model

The SEIR(susceptible, exposed, infectious, and recovered) epidemiologic model partitions a population into four groups or compartments. Starting from an initial state individuals move among the four compartments at rates described by the following equations:

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dE}{dt} = \beta SI - \sigma E$$
$$\frac{dI}{dt} = \sigma E - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

$$\beta = \mathrm{R}_0 \gamma / N$$
$$\gamma = 1/\tau$$

where

S	is the susceptible population
E	is the exposed population
I	is the infected population
R	is the recovered population
R_0	is the transmissibility parameter
τ	is the mean duration of the infection
N	is the total population
t	is time

Solving this system of four differential equations results in four functions that describe the number of people in each partition of the population over time; $S = S(t)$, $E = E(t)$, $I = I(t)$, and $R = R(t)$.

In order to apply this model it is important to understand the assumptions that it makes. One limitation is that it only applies to variables and parameters that are smooth, continuous functions of time. Another assumption of this SEIR model is that its differential equations do not include stochastic error terms. The addition of stochastic error terms may not apply to the dynamics describing the underlying spread of an epidemic, and they would greatly complicate the analysis of this model. However, the consideration of uncertainties is crucial to any study based on the SEIR model.

Another characteristic of this model the total population remains constant because for each term in one of the rate equations there is a corresponding term in another rate equation with an opposite sign. For example the outflow of people from the susceptible compartment, $-\beta SI$, is matched by a correspoding inflow term in the exposed compartment, βSI.

The Method

Because the SEIR model represents the rate of change in the dependent variables as a nonlinear function of S and I it is not possible to derive an analytic solution for the dependent variables over time. It is possible to solve the SEIR differential equations using a numerical approximation though. For cases where the initial conditions and model parameters lead to a well-conditioned model there are many numerical approximations available which can integrate this model over time to simulate the trajectory of the dependent variables. In some cases; however, variations in the magnitude of the four differential equations in the SEIR model cause a simple approximation to the dependent variables like the following to become inaccurate:

$$
\begin{aligned}
S_{t+1} &\approx -\beta S_t I_t \Delta t \\
E_{t+1} &\approx (\beta S_t I_t - \sigma E_t)\Delta t \\
I_{t+1} &\approx (\sigma E_t - \gamma I_t)\Delta t \\
R_{t+1} &\approx \gamma I_t \Delta t
\end{aligned}
$$

where Δt is the time step. When the SEIR model is expressed in the following vector form:

$$
\frac{d\mathbf{v}}{dt} = \mathbf{f}(\mathbf{v}, t) \qquad \text{where} \quad \mathbf{v} = \begin{pmatrix} S \\ E \\ I \\ R \end{pmatrix}
$$

the Jacobian matrix, $\mathbf{J}(t)$, can be calculated as

$$
\mathbf{J}(t) = \frac{\partial \mathbf{f}}{\partial \mathbf{v}}
$$

If the condition number of $\mathbf{J}(t)$ becomes large, errors in the numerical approximations to the integration can become large and compromise simmulation results. PROC TMODEL is well suited for differential equation models like this one because it can control the amount of error introduced by the numerical integation process even when models becomes ill-conditioned.

Complications Exposed By Simulations

Initial values of the dependent variables must be provided in addition to values of the model parameters to simulate the SEIR equations. In the following simulation the initial values; $S_0 = S(t = 0)$, $E_0 = E(t = 0)$, $I_0 = I(t = 0)$, and $R_0 = R(t = 0)$; are specified in the TIMEPTS data set. Also, a parameter covariance matrix, ESTDATA=, is used in the SOLVE statement to specify how parameters are varied in the Monte Carlo simulations.

```
proc tmodel data=timepts;
   endo s &N e 0 v 1 r 0;
   /* parameters of interest */
   parms R0 &Rho0 tau &tau;

   /* fixed values */
   control N  &N
           sigma  &sigma;

   /* coefficient parameterizations */
   gamma = 1/tau;
   beta = R0*gamma/N;

   /* Differential equations */
   dert.s = - beta*s*v;
   dert.e = beta*s*v - sigma*e;
   dert.v = sigma*e - gamma*v;
   dert.r = gamma*v;

   cases = v + r;

   /* monte carlo simulation; r0 */
   solve cases / time=date outpredict out=mcsimr0(rename=v=i) seed=1
                 random=5 sdata=s estdata=r0cov;
   outvars R0 tau;
   /* monte carlo simulation: r0 and tau */
   solve cases / time=date outpredict out=mcsimr0t(rename=v=i) seed=1
                 random=10 sdata=s estdata=r0tcov;
   outvars R0 tau;
quit;
```

Figure 0.1 shows how the peak and duration of the infection are impacted by different values of R_0. This analysis confirms the intuitive understanding that decreasing R_0 lowers the peak and delays the resolution of the epidemic.

Figure 0.1 Monte Carlo Simulations of Infection Trajectories for R_0

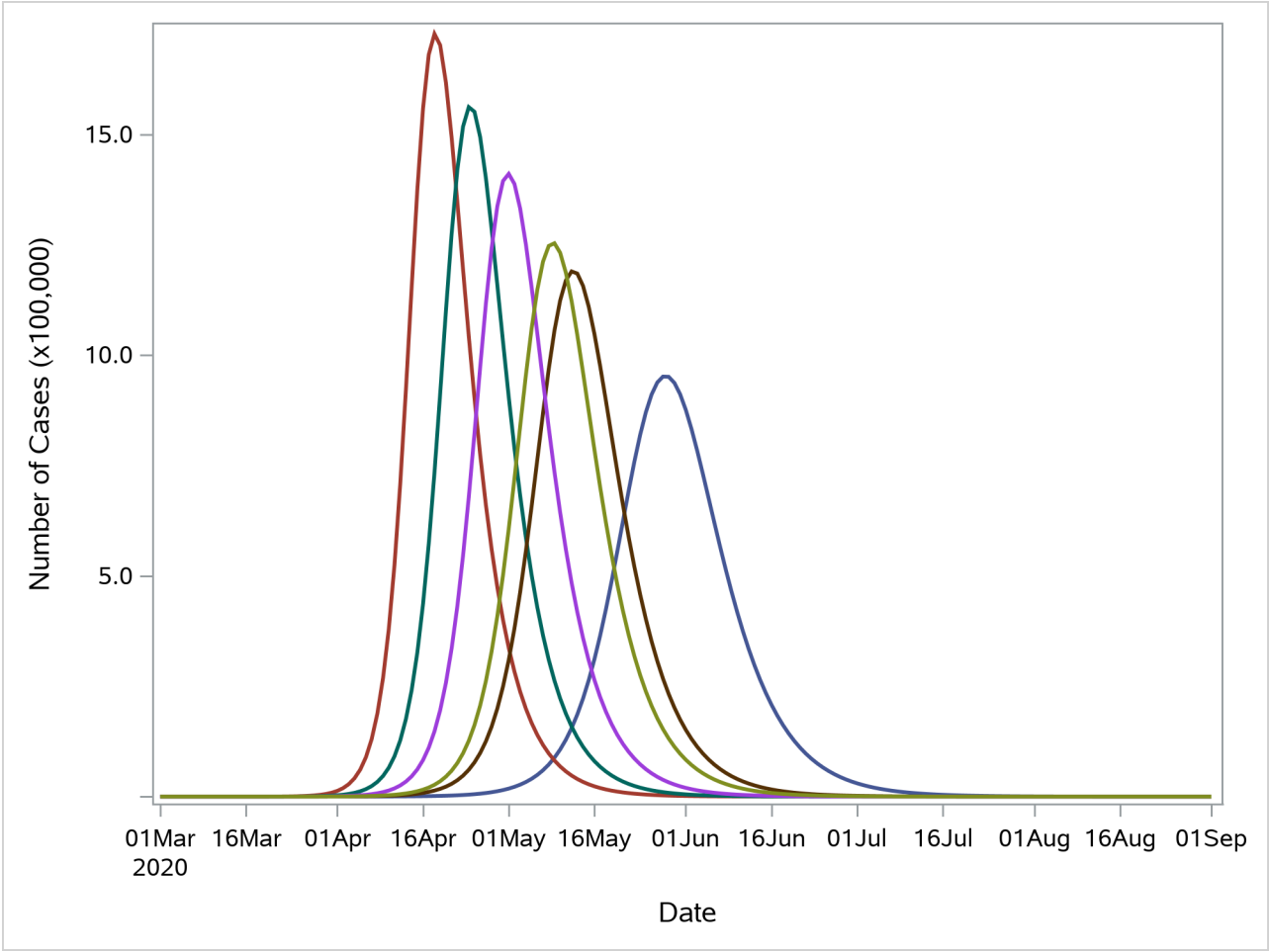
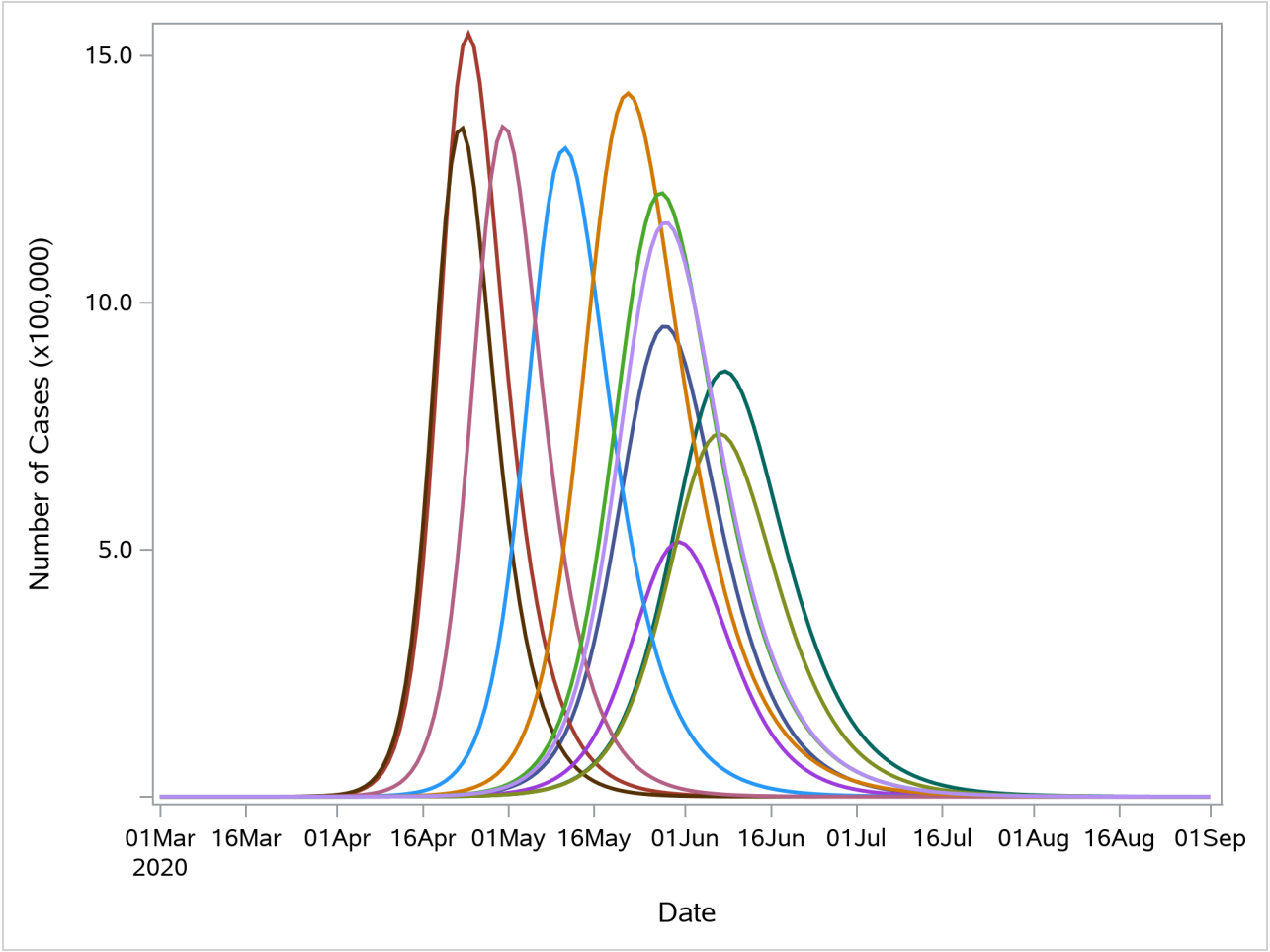


Figure 0.2 shows results of a Monte Carlo simulation that varies both R_0 and τ. While Figure 0.1 is easily understood, the intensities and durations of the curves in Figure 0.2 present a much more complicated picture of the possible resolutions of the epidemic when uncertainties in two parameters are considered. When more parameters are varied in the simulations the sensitivities of the SEIR model becomes even murkier.

Figure 0.2 Monte Carlo Simulations of Infection Trajectories for R_0 and τ



Trying to Reduce Uncertainties Through Estimation

When the full possible domain of unknown parameters' values are considered in Monte Carlo simulations, the predicted range of infection trajectories is too large to be of much use. One way to reduce the uncertainty in parameter values is to estimate them using observed infection data. Because different geographic populations have unique characteristics both for transmissibility, and the medical treatment given to infected individuals, a population's own observed infection data is useful in estimating $\hat{\mathrm{R}}_0$, $\hat{\tau}$ and other parameters. These population estimates may then inform the choice of parameters' values used in subequest Monte Carlo simulations.

The following code estimates I_0 and R_0 for North Carolina and Ohio:

```
proc tmodel outmodel=seirmod;
   /* Parameters of interest */
   parms R0 &Rho0 i0 1;
   bounds 1 <= R0 <= 13;
```

```
    /* fixed values */
    control N  &N
            sigma  &sigma
            tau &tau;

    /* coefficient parameterizations */
    gamma = 1/tau;
    beta = R0*gamma/N;

    /* Differential equations */
    dert.s = - beta*s*v;
    dert.e = beta*s*v - sigma*e;
    dert.v = sigma*e - gamma*v;
    dert.r = gamma*v;
    cases = v + r;

    outvars s v e r;

    /* Fit the NC data */
    fit cases init=(s=&N v=i0 r=0 e=0) / time=date dynamic outpredict outactual
                out=ncpred(rename=v=i) covout outest=ncprmcov optimizer=ormp(opttol=1e-5)
                ltebound=1e-10 data=jhuus(where=(Province_State="North Carolina"
                and cases>0 and date < &enddate));
    /* Fit the OH data */
    fit cases init=(s=&N v=i0 r=0 e=0) parms=(R0 &Rho0 i0 1) / time=date dynamic
                outpredict outactual out=ohpred(rename=v=i) covout outest=ohprmcov
                outs=cov_error optimizer=ormp(opttol=1e-5) ltebound=1e-10 data=jhuus(where=
                (Province_State="Ohio" and cases>0 and date < &enddate));
quit;
```

The parameter estimates computed by the PROC TMODEL fits to aggregated county data published by Johns Hopkins University are shown in Table 0.3, and Table 0.4. Figure 0.5 compares the data to TMODEL's SEIR predictions.

Figure 0.3 North Carolina estimates of R_0 and I_0

The TMODEL Procedure

Nonlinear OLS Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
R0	2.737459	0.0522	52.41	<.0001
i0	2.677113	0.3183	8.41	<.0001

Figure 0.4 Ohio estimates of R_0 and I_0

Nonlinear OLS Estimates (Not Converged)

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
R0	2.917217	0.0587	49.73	<.0001
i0	12.58116	1.1132	11.30	<.0001

Figure 0.5 Estimation of R_0 and I_0

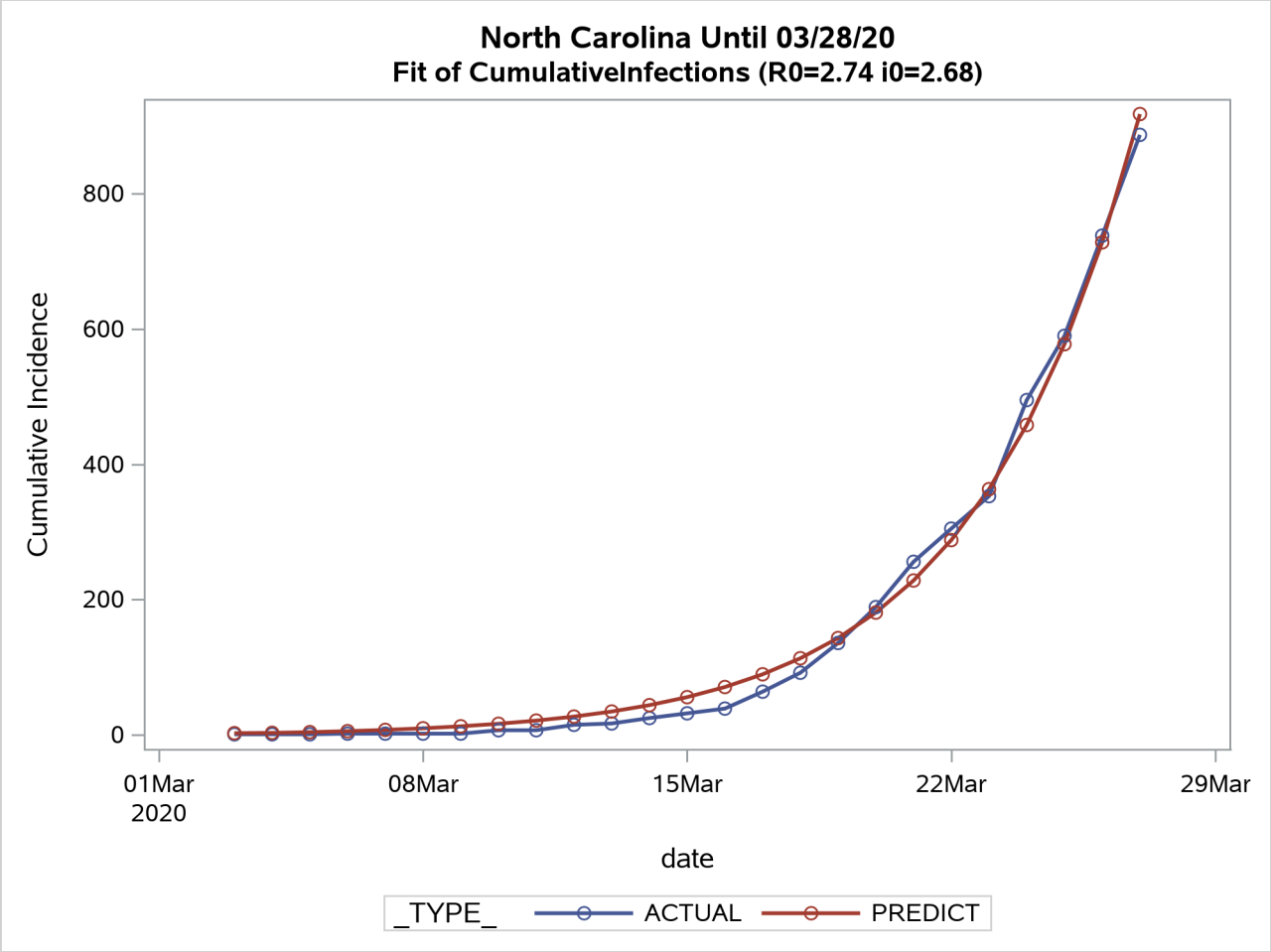
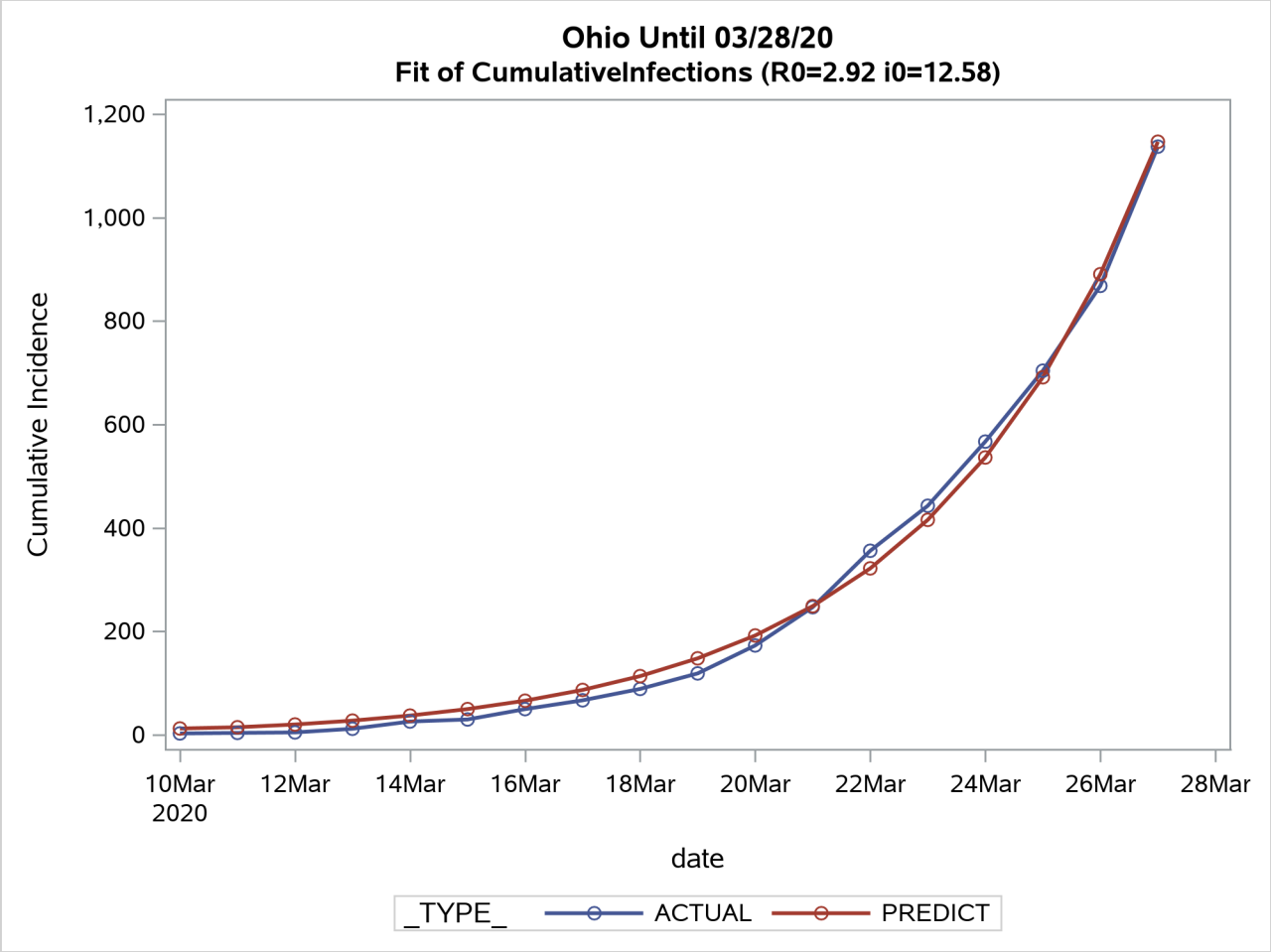


Figure 0.5 *continued*



Adding an Intervention Term

The SEIR model fits these states' data well up until the last week of March. Around that time the growth in the infection rate appears to slow due to factors that are not included in the model. It is likely that social distancing measures and other mitigation efforts are impacting the spread of the disease. To gain further insight on the magnitude of the iterventions, an adjustment to R_0 can be introduced to the original model that allows the R_0 parameter to change over the course of the epidemic. The following alternative parameterization of β in the SEIR model achieves this goal:

$$\beta = \beta(t) = (\mathrm{R}_0 + \mathrm{R}_i \Phi(t - t_i)) \frac{\gamma}{N}$$

where R_i is the achieved change in R_0, t_i is the mean time the intervention takes effect, and $\Phi(x)$ is the standard normal cumulative distribution function.

In order to fit the SEIR model including the intervention adjustment data for the population should include sufficient data after the time when the intervention was achieved. The following PROC TMODEL estimations

use data for Washigton state and Italy since it is anticipated their data represent populations in the US and Europe that have the most data to support this analysis:

```
proc tmodel outmodel=seirmod;
   /* Parameters of interest */
   parms R0 &Rho0 i0 1 Ri -1 di &idate;
   bounds 1 <= R0 <= 13;
   restrict R0 + Ri > 0;

   /* fixed values */
   control sigma  &sigma
           tau &tau;

   /* coefficient parameterizations */
   gamma = 1/tau;
   step = cdf('normal',date,di,1);
   beta = (R0 + step*Ri)*gamma/N;

   /* Differential equations */
   dert.s = - beta*s*v;
   dert.e = beta*s*v - sigma*e;
   dert.v = sigma*e - gamma*v;
   dert.r = gamma*v;
   cases = v + r;

   outvars s v e r;

   /* Fit the Italy data */
   fit cases init=(s=&Nit v=i0 r=0 e=0) / time=date dynamic outpredict outactual
               out=itpred(rename=v=i) covout outest=itprmcov optimizer=ormp(opttol=1e-5)
               ltebound=1e-10 data=italy(where=(cases>0));
   /* Fit the Washington data */
   fit cases init=(s=&Nwa v=i0 r=0 e=0) parms=(R0 &Rho0 i0 1 Ri -1 di &idate) / time=date
               dynamic outpredict outactual out=wapred(rename=v=i) covout outest=waprmcov
               outs=cov_error optimizer=ormp(opttol=1e-5) ltebound=1e-10
               data=washington(where=(cases>1));
quit;
```

Table 0.6 and Table 0.7 show the results of adding am intervention term to the exposure rate, and Figure 0.8 show a plot of how the new model's predictions represent the observed case data.

Figure 0.6 *continued*

Figure 0.6 Washington estimates of intervention model

The TMODEL Procedure

Nonlinear OLS Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
R0	2.09994	0.0254	82.59	<.0001
i0	14.33301	1.9341	7.41	<.0001
Ri	-1.08396	0.0254	-42.63	<.0001
di	21994.02	0.1825	120530	<.0001
Restrict0	0.002529	72.9878	0.00	1.0000

Figure 0.7 Italy estimates of intervention model

Nonlinear OLS Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
R0	2.235483	0.1431	15.63	<.0001
i0	31.19957	8.1032	3.85	0.0005
Ri	-0.7471	0.1407	-5.31	<.0001
di	21994.19	1.3697	16058.0	<.0001

Figure 0.8 Intervention model comparisons to infection incidence

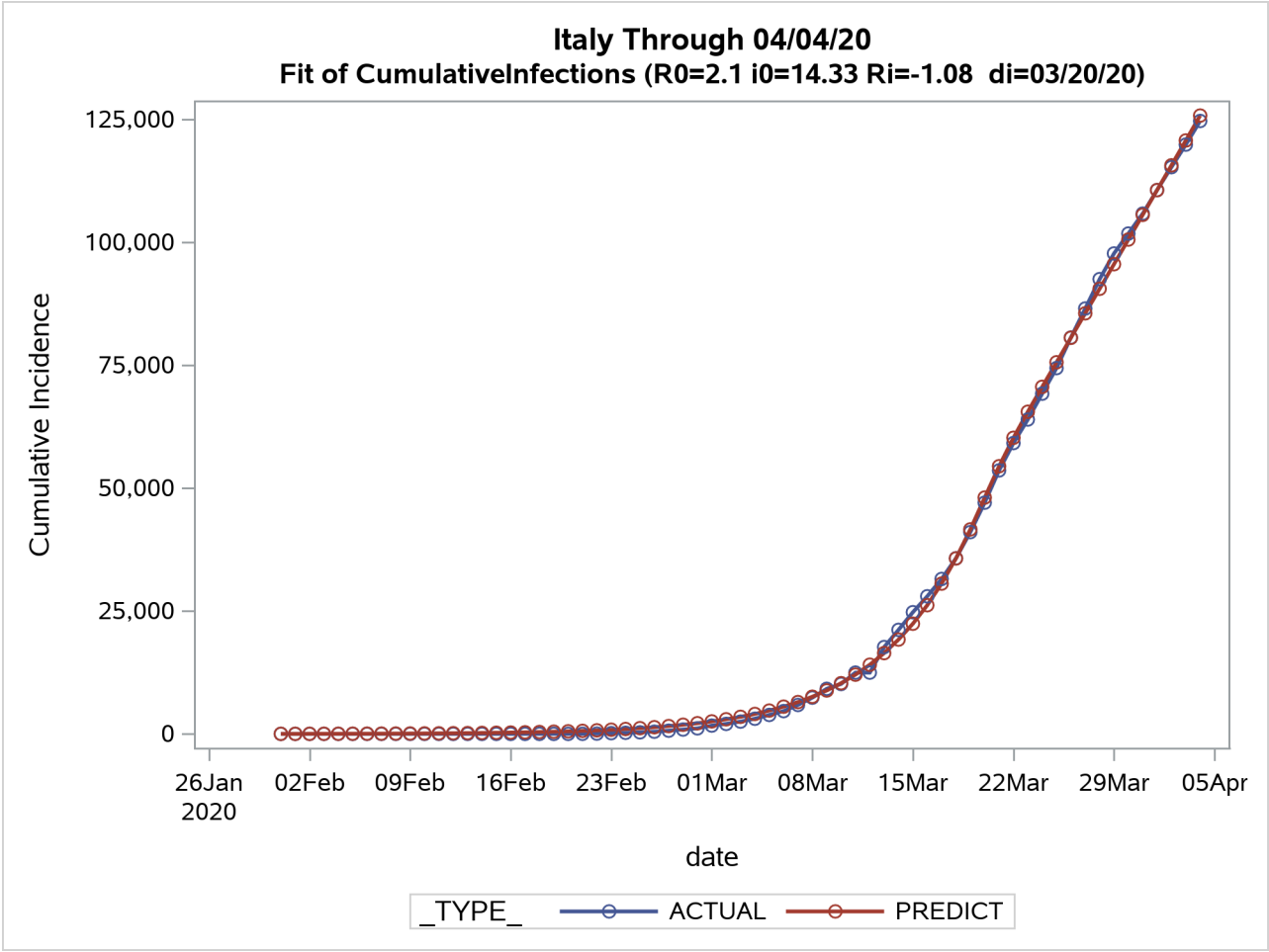
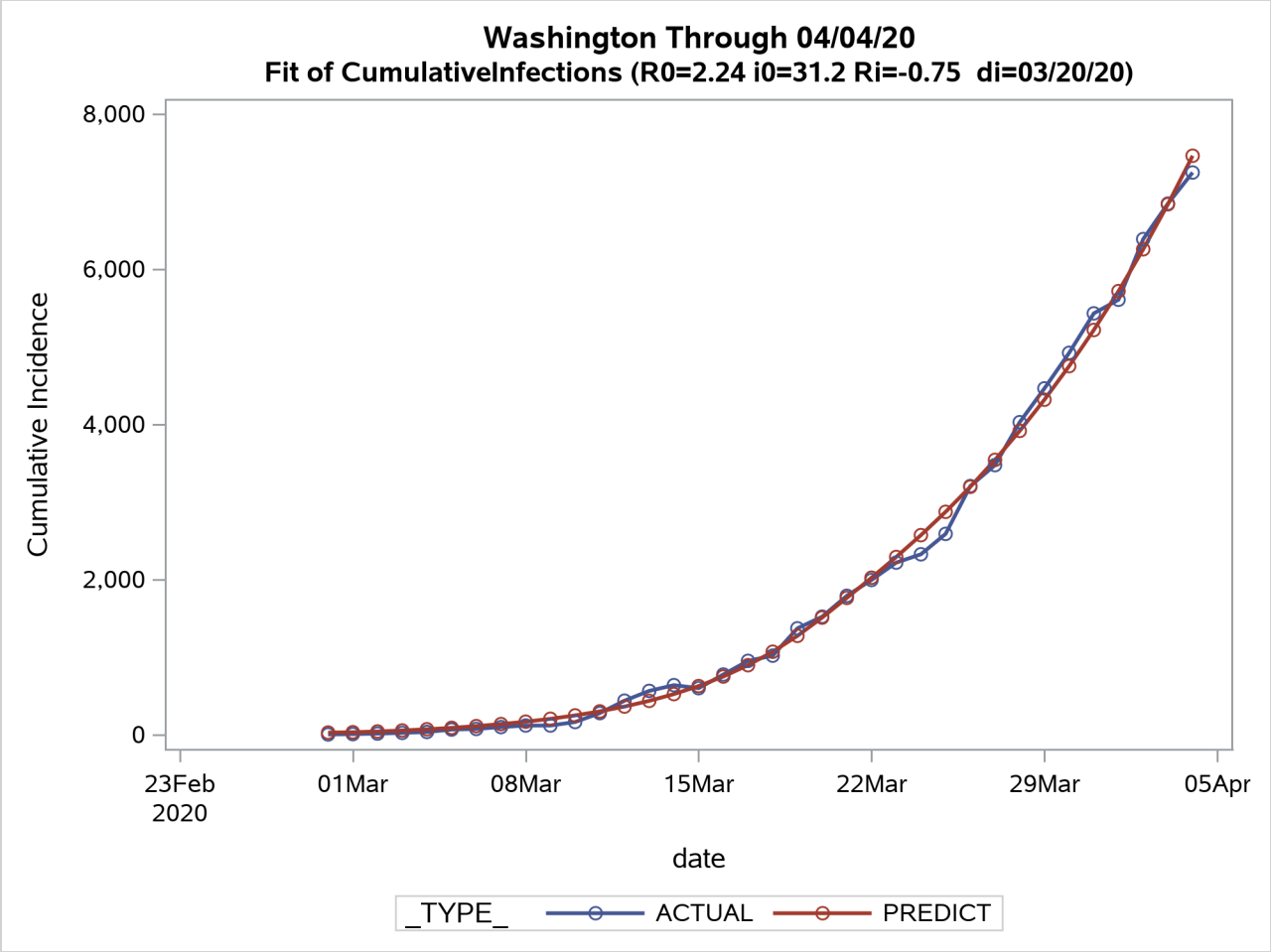


Figure 0.8 *continued*



Parameter Sensitivities and the Perils of Forecasting

The SEIR model with a time dependent intervention term accurately models COVID-19 incidence in all geographic regions that have been analyzed so far. The fidelity of the model to various populations with no more than four parameters is impressive, but it also hints at a shortcoming of the model. Future predicted infection incidence trajectories can vary widely due to small changes in the parameters. The same qualities of the model that allow it to fit data well result in forecasts that are extremely sensitive to small variations in the parameters. One indication of the parameter sensitivities is the small standard errors reported in the estimation results. See Table 0.6 and Table 0.7 for example.

Many analyses of parameter sensitivities are possible for the SEIR model. A version of the model with an intervention term that anticipates the resolution of the epidemic is

$$\beta = \beta(t) = \mathrm{R}_0(1 - \Phi(t - t_i, \sigma_t^2))\frac{\gamma}{N}$$

where $\Phi(x, \sigma_x^2)$ is the cumulative normal distribution with mean x and variance σ_x^2. This intervention term is an optimistic version of the resolution of the epidemic because it allows for the mitigation efforts to reduce the R_0 value to zero as an alternative to the longer term, classical resolution of the epidemic due the population being saturated with immune individuals.

```
title 'SEIR Fit Model';
proc tmodel outmodel=seirmod;
   /* Parameters of interest */
   parms R0 &Rho0 i0 1 /*Ri -1*/ di &idate dstd=10;
   bounds 1 <= R0 <= 3;
   /*restrict R0 + Ri > 0; */

   /* fixed values */
   control sigma  &sigma
           tau &tau;

   /* coefficient parameterizations */
   gamma = 1/tau;
   step = cdf('normal',date,di,dstd);
   beta = (R0 - step*R0)*gamma/N;

   /* Differential equations */
   dert.s = - beta*s*v;
   dert.e = beta*s*v - sigma*e;
   dert.v = sigma*e - gamma*v;
   dert.r = gamma*v;
   cases = v + r;

   outvars s v e r;

   /* Fit the Italy data */
   fit cases init=(s=&Nit v=i0 r=0 e=0) / time=date dynamic outpredict outactual out=itpred(r
               covout outest=itprmcov optimizer=ormp(opttol=1e-5) ltebound=1e-10
               data=italy(where=(cases>0));
quit;
```

The forecasts for Monte Carlo simulations of this optimistic intervention model are shown using two different y-axis scales in Figure 0.9. This simulation indicates a large variance in the number of individuals that would be infected based on how quickly effective mitigation practices can take full effect. However, in all simulations the epidemic resolves very quickly once full mitigation is achieved.

Figure 0.9 Forecast of infection incidence for Italy

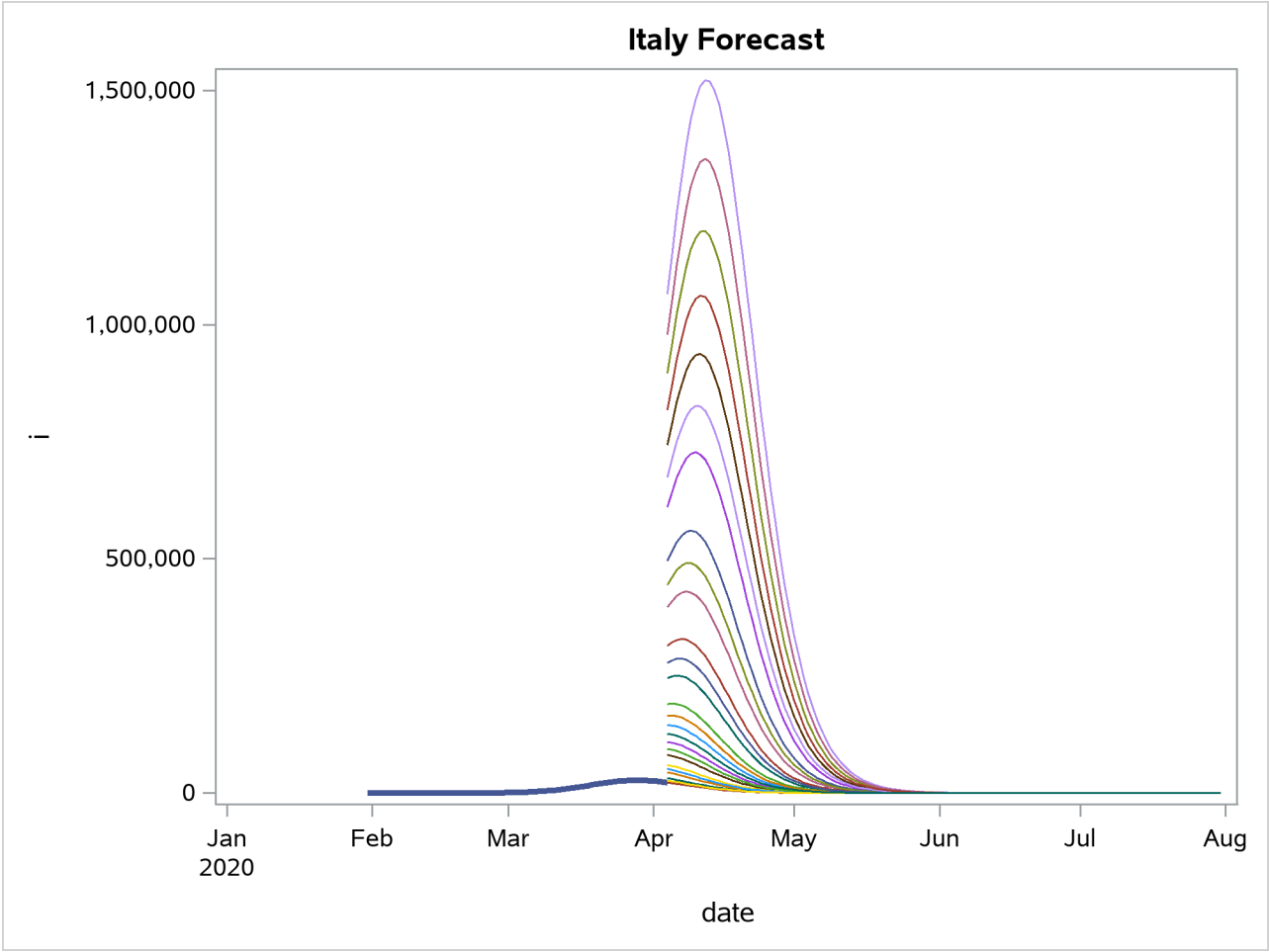


Figure 0.9 *continued*