

# pomp data

Andrew Tredennick

4/5/2020

**pomp** wants time-indexed case reports when using **accumvars** that reset to 0 at each observation timestep.

That's the punchline; let's prove it.

## Evidence from examples

The codes for King's **dacca**, **ebola**, and **sir** examples all have time-indexed case reports (new cases or some index thereof) as the data. None of these use cumulative case counts as the data. Likewise, these all use the **accumvar** definition for the variables passed to the likelihood evaluator.

Here's the **dmeasure** snippet for the **sir** example explored below:

```
dmeasure = Csnippet(  
  " double mean, sd;  
    double f;  
    mean = cases*rho;  
    sd = sqrt(cases*rho*(1-rho));  
    if (reports > 0) {  
      f = pnorm(reports+0.5,mean,sd,1,0)-pnorm(reports-0.5,mean,sd,1,0);  
    } else {  
      f = pnorm(reports+0.5,mean,sd,1,0);  
    }  
    lik = (give_log) ? log(f) : f;"  
)
```

And here's the **accumvar** setting for the simulated **cases** variable that is evaluated relative to the **reports** data in the likelihood function.

```
accumvars = c("cases")
```

As we'll see below, both **cases** (sims) and **reports** (data) need to both be new cases, not cumulative cases.

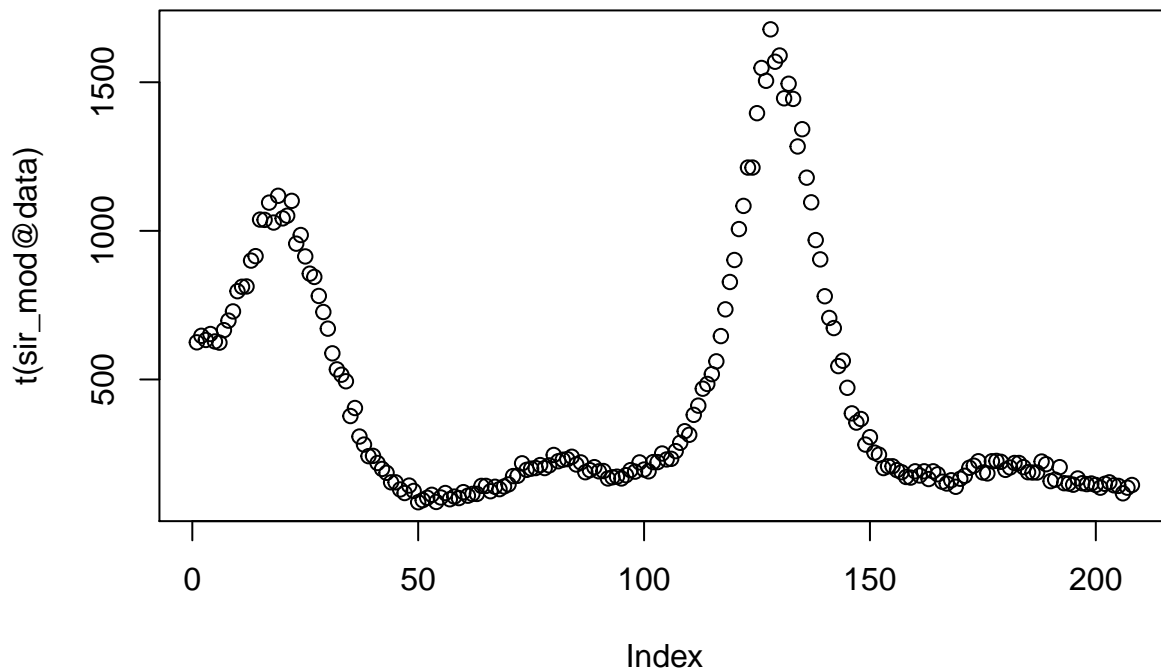
## Evidence from quick simulation

Here I use the **pomp::sir** model and simulate some data. First, I use **pfilter** to compute the likelihood of the model fit to time-indexed reports of new cases. Second, I use the same procedure to the likelihood of the model fit to cumulative reports.

### Model set up

Note that the data supplied with the **pomp** object are new reports (they go up and down)

```
sir_mod <- sir() # load the pomp object  
  
plot(t(sir_mod@data)) # data that is shipped with pomp
```



### Simulate data

```
# Simulate a data set
sim_data <- simulate(sir_mod, nsim = 1, format="data.frame",
                    include.data = FALSE)

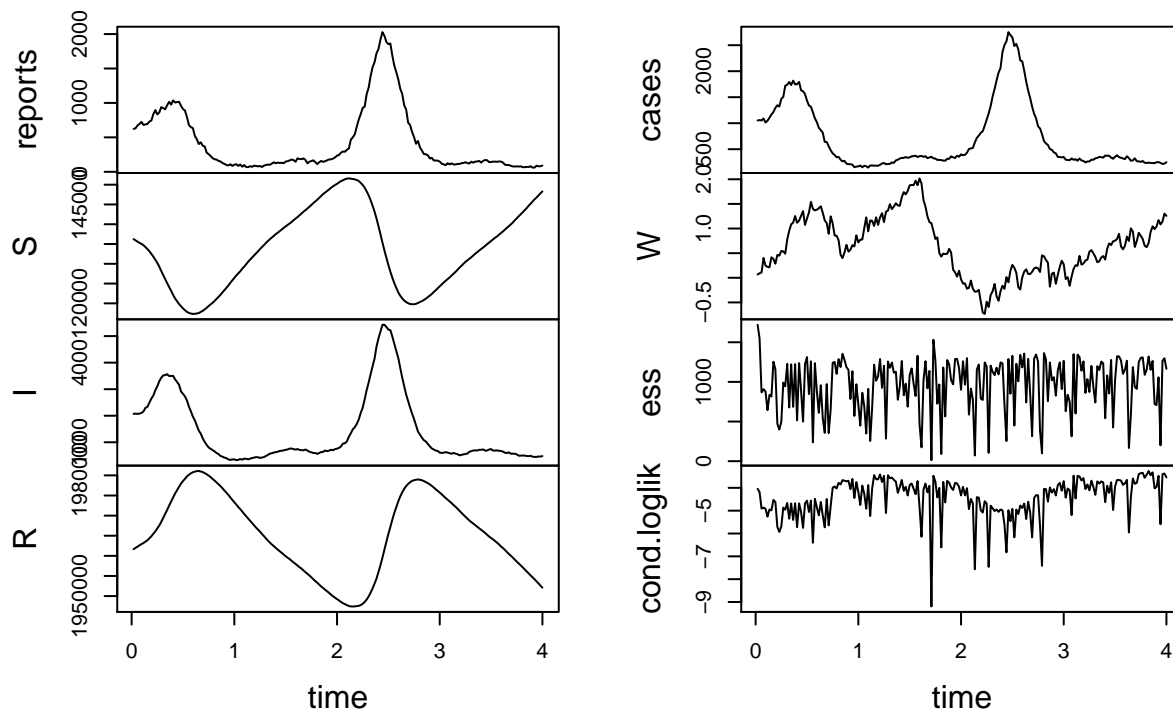
sir_mod@data[1, ] <- sim_data$reports # replace data

coef(sir_mod) # print the coefficients

##      gamma      mu      iota    beta1    beta2    beta3  beta_sd      rho
## 2.60e+01 2.00e-02 1.00e-02 4.00e+02 4.80e+02 3.20e+02 1.00e-03 6.00e-01
##      pop      S_0      I_0      R_0
## 2.10e+06 6.50e-02 1.00e-03 9.34e-01
```

### Fit to new reports

```
pf <- pfilter(sir_mod, Np = 2000)
plot(pf)
```



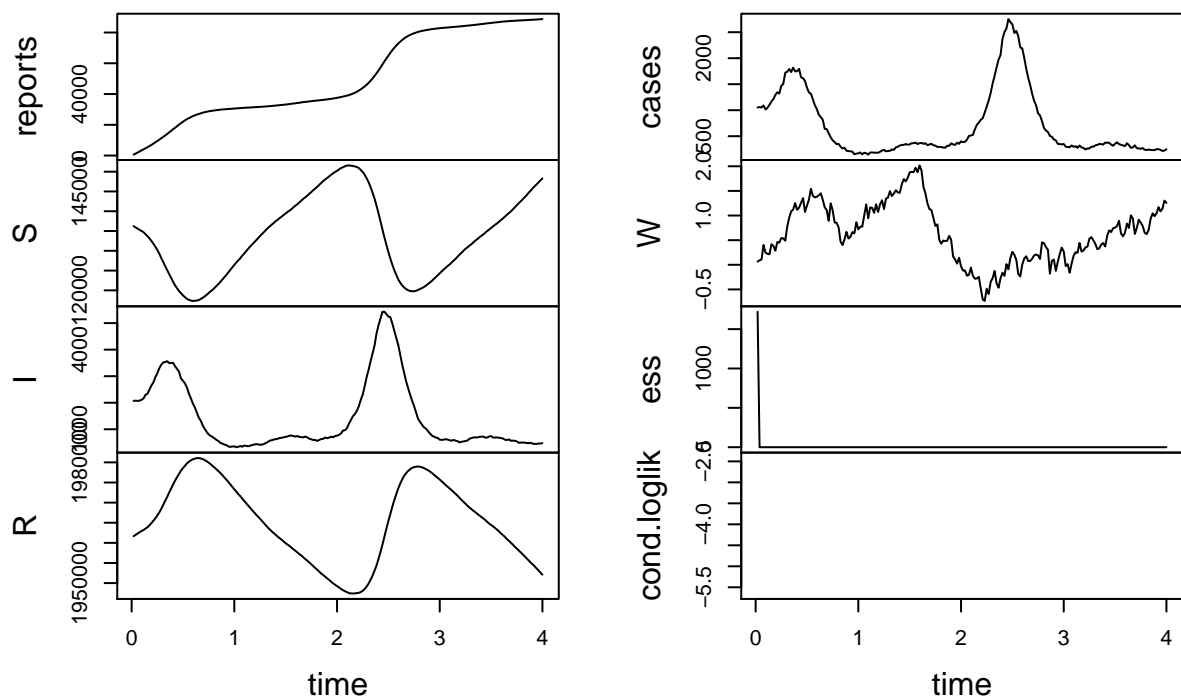
## Fit to cumulative reports

Note that there are many, many filtering failures; that the “reports” (data) and cases (sims) do not match up; and that the likelihood cannot even be evaluated.

```
sir_mod@data[1, ] <- cumsum(sim_data$reports) # replace data with cumulatives
pf2 <- pfilter(sir_mod, Np = 2000)
```

```
## Warning: in 'pfilter': 207 filtering failures occurred.
```

```
plot(pf2)
```



## Conclusion

We need to use daily reports of new cases as the data in our fitting routines. This means using `diff(covid_ga_data$cases)` as the data.