Six Lessons Software Estimation Can Learn from Covid-19 Forecasting

By Steve McConnell
My personal role in Covid-19 Forecasting
Experts tend to use simple estimation strategies, even when their level of expertise in the subject being estimated is high.
— Steve McConnell, *Software Estimation*
Observed and forecasted new and total reported COVID-19 deaths as of September 28, 2020.

Interpretation of Forecasts of COVID-19 Deaths

Leading edge
- The leading edge of a forecast is the earliest date when a forecast is made.

Subscribers
- Subscribers to the forecast are notified of the new forecast via email.

Corrections
- Corrections to the forecast are made when new data is available.

References
- References for the forecast are listed at the bottom of the page.
National Forecast

To receive email updates about COVID-19, enter your email address:
National Forecast

New Weekly Deaths

[Graph showing weekly deaths with various forecasts and bands for 95% prediction intervals]
The Basics:

- Estimate
- Forecast
- Speculation
A forecast is a future-based estimate

- I can “estimate” the number of fish in a pond
- I can either “forecast” or “estimate” how many iterations my next release will require
Forecasting is not the same as speculation, but we are seeing them treated interchangeably.
The Basis of
Forecasting/Estimation

“What We Know”
CDC Covid Forecasting
CDC Covid-19 Forecast Process

CDC work is overseen by a team at University of Massachusetts Amherst.

Specific organization is Reich Lab based in the Department of Biostatistics and Epidemiology.

Modeling groups submit forecasts to the CDC.

Forecasts that meet certain criteria are combined into the “Ensemble” model, which is the forecast model of record for the CDC.

#CovidComplete by Steve McConnell
CDC Forecasting
What Each Team Contributes Weekly

- ~30 modeling teams per week since July 2020
- Forecasts for US National, 50 US states + DC (at least)
- 1, 2, 3, and 4 week forecast horizons
  - Incremental forecasts (1 week at a time)
  - Cumulative forecasts (total for pandemic to date)
- Forecasts for these quantiles:
  - 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.975, 0.99
- Total of about 10,000 forecasts per week, per team
CDC Forecasting
What I Analyze

- Forecasts for US National, 50 US states + DC
- 1, 2, 3, and 4 week forecast horizons
  - Incremental forecasts, cumulative from forecast date
- Forecast types:
  - Point forecast (0.50 quantile)
  - 50% Prediction Interval (PI) forecast (0.25 to 0.75 quantile range)
  - 95% Prediction Interval (PI) forecast (0.025 to 0.975 quantile range)
- Subset = ~625 forecasts per week, per team
CDC Forecasting

Total Data Set of Forecasts vs. Actuals

- ~6,500 forecast sets
- ~625 forecasts per set
- Total of ~4 million forecasts vs. actuals to date
- Overall, an incredibly rich data set in which to explore topics like different estimation methods, accuracy at different time horizons (i.e., cone of uncertainty), how well models deal with uncertainty, etc.
CDC Forecasting Context

Forecasts and actuals are all public information

- Detailed Forecasts: https://github.com/reichlab/covid19-forecast-hub
- Actuals: https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
- Evaluation data: https://github.com/stevemcconnell/covid19-forecast-evaluations
- Evaluation graphs: https://stevemcconnell.com/covidcomplete
Covid Complete’s Forecasting Approach
Timing forms the basis of forecasting (for my method)

We can do this with terrific accuracy at the national level, and with fair accuracy at the state level.

And because trends tend to continue, we can project forward for another 7-14 days.

We can take the test number, which we know on date X.

We can apply what we’ve learned about the ratio of positive tests to deaths and the lag time.

We can forecast the deaths number, 14-21 days ahead.

#CovidComplete by Steve McConnell

stevemcconnell.com/covidcomplete
But is it working in practice? I.e.,

How is CovidComplete performing?
How is CovidComplete Performing?

US National Point Fatality Forecasts from Aug 24 to May 24, 2021
Forecasts ending May 29, 2021

For national forecasts, CovidComplete has been among the top 5 models over the long term.
Use Historical Data
Most Common Forecast Method: SEIR

- SEIR = Susceptible, Exposed, Infected, Recovered
- This is the established epidemiological model
- Forecast results from these models were mediocre for many months but have improved recently
  - Mistaken old assumptions must be removed from the models, e.g., from flu forecast models
  - Mistaken new assumptions take time to invalidate and then remove
Other Methods

- Massive data sets, e.g., 500 million records
- Esoteric data, e.g., use of mobility data from cell phone records
- Machine learning
- Artificial intelligence
- Bayesian analysis, Monte Carlo simulations, etc.
My forecasting method
The essence of my forecasting method:

- Count positive tests
- Calculate current CFR (death rate)
- Multiply tests by current CFR
The Best Approach So Far

- Simple models based on positive tests and the lag from positive tests to deaths
- A few of the more accurate models use this simple approach, including CovidComplete
Software Lesson Learned

Use Historical Data

- Your historical data accounts for most or all factors that affect your productivity, and it is not subject to estimation bias
  - E.g., velocity
- Put your effort into being sure your historical data is accurate and means what you think it means
Keep “Control Knobs” to a Minimum
Groups use Various Methods

- Massive data sets, e.g., 500 million records
- Esoteric data, e.g., use of mobility data from cell phone records
- Machine learning
- AI
- Bayesian analysis, Monte Carlo simulations, etc.
- Pre-existing infectious diseases models

Model descriptions are available from the [CDC website](https://www.cdc.gov) (many just link to github)
One of the most enduring and useful conclusions from research on forecasting is that simple methods are generally as accurate as complex methods.

— J. Scott Armstrong, *Principles of Forecasting*
Keep “Control Knobs” to a Minimum

- Favor simple models where you understand the data you’re using
- In general, this means proxy-based estimation
- Examples include story points, stories, features per week/month/sprint, etc.
Accuracy First, Precision Second
Accuracy vs. Precision

- **Accuracy** = Did you hit the mark?
  - For prediction intervals, this refers to whether the PI captured the intended percentage of actuals, e.g., a 95% PI should capture about 95% of actuals

- **Precision** = How exact was your forecast?
  - In engineering, precision is normally the number of significant digits used to represent a quantity
  - In forecasting, precision is the width of the forecast range
  - The Ensemble team refers to this as “sharpness”
State 95% Prediction Range Fatality Forecasts for the 4 Week Period Ending 2021-1-23

Forecast date: 2020-12-28

CovidComplete

Source: covidcomplete.org
Software Lesson Learned

Accuracy First, Precision Second

- Avoid presenting estimates that mismatch accuracy and precision, e.g., avoid “This release will take 4 sprints, plus or minus 3 sprints”
Understand the Variability in Your Data
Data Issue #1: Corrections
Spikes (Corrections) in State Level Data

Daily Positive Tests and Deaths in WA as of 5/30/21

Source: stevemcconnell.ccm/covidcomplete
Corrections in the Past 7 Days

- Large numbers of states did not update on Memorial Day
- WA added ~3000 backlogged cases from May 26-29
- WI added 41 backlog deaths May 27
- NM added 100 backlog deaths May 24
- TX added 628 backlogged cases May 24
- OK added ~350 deaths May 26
- MD added 583 backlogged deaths May 27
- LA added 3857 backlogged cases May 27
- WY added 100+ cases May 20
Data Issue #2: Sundays
Issue with Daily Data – Sundays!
Variation in Data Reporting by Day-of-Week

Covid-19 Death Reporting's Day-of-Week Deviation from Trend

- Sun: -45%
- Mon: -48%
- Tue: 16%
- Wed: 34%
- Thu: 24%
- Fri: 16%
- Sat: 3%

Source: stevemcconnell.com/covidcomplete
Data Issue #3: Holidays
Covid-19 Death Reporting's Deviation from Trend
Adjusted for normal day-of-week variation

Labor Day

9/5  9/6  9/7  9/8  9/9  9/10  9/11  9/12  9/13  9/14  9/15  9/16  9/17  9/18

-100.7% -63.5%  -0.3%  13.7%  8.1%  -4.1%  -10.1%  -0.8%  10.7%  9.7%  -10.5%  0.0%
Holidays

Thanksgiving Weekend
(4-day weekend, apparently)

Covid-19 Death Reporting's Deviation from Trend
Adjusted for normal day-of-week variation
Holidays

Covid-19 Death Reporting's Deviation from Trend
Adjusted for normal day-of-week variation

Christmas
Wrongly classified deaths and subsequent corrections
  • Adding deaths “attributed to” CV
  • Subtracting deaths classified as “with CV” that were not “from CV”
Wrongly classified tests (PCR vs. antibody) and subsequent corrections
Data initially reported on wrong date, and later corrected
Deaths initially reported in one state and later changed to another
Data Issues: “Common Cause” Variations

- Weekly reporting cycle
- Holiday under- and over-reporting (which is a mix of common cause and special cause)
Software Lesson Learned

Understand the Variability in Your Data

- Variations in velocity
- People’s personal schedules
- People’s allocation to your project
- Holidays, sick days, training days, etc.
Minimize Differences Between Reporting and Ground Truth
We’re really working with 3 entities, not 2:

- Forecast
- Ground Truth
This is more about reporting dates, more than actual dates.
We’re really working with 3 entities, not 2:

- Forecast
- Ground Truth
- **Reporting** of the Ground Truth
Minimize differences between reporting and ground truth

- Do you actually finish stories each sprint. If not, is your velocity really “ground truth” or is it really an estimate of ground truth?

- Do you change point values of stories after completing them (“This was more like a ‘5’ than a ‘3’”)?

- Did you really meet the sprint goal, or was the sprint goal redefined so that you could meet it?

- Are you accounting for weekends and holidays in your estimates and plans?
Close the Loop
CovidComplete Forecasts for Covid-19 Deaths in the US, forecast date of 5/8/21

This is a forecast of record submitted to the CDC

2-Week Forecasts
Low  
Nominal  7,250  
High  10,250

1 week error: -6.2%
2 week error: -5.3%

Source: stevemcconnell.com/covidcomplete
Remember this chart?

US National Point Fatality Forecasts from Aug 24 to May 24, 2021
Forecasts ending May 29, 2021

- Middle 50%
- Median Forecast Bias
- Worst High and Low Forecasts

#CovidComplete by Steve McConnell

stevemcconnell.com/covidcomplete
Remember this Chart?

These combinations of accuracy and precision are rarely possible (<0.5% of all forecasts).

source: covidcomplete.org
This is what it looks like now

These combinations of accuracy and precision are rarely possible (<0.5% of all forecasts)

source: covidcomplete.org
Look at the improvement! (Past)
Look at the improvement! (Present)

These combinations of accuracy and precision are rarely possible (<0.5% of all forecasts)


*Forecast fewer than 50 states+DC

source: covidcomplete.org
Software Lesson Learned

Close the Loop

- Compare estimates to actuals—see how you did!
- When there are gaps, be sure you understand the reasons
- Iterate your estimation process based on this error feedback
- Inspect and adapt: this is arguably the most important point in this talk
Summary
Summary

- Use historical data
- Keep “control knobs” to a minimum
- Accuracy first, precision second
- Understand the variability in your data
- Minimize differences between reporting and ground truth
- Close the loop
Background on Much of This is in My Estimation Book
Useful Links

Covid-19 Resources

• SteveM’s Covid-19 Data Center
  https://www.stevemcconnell.com/covid

• SteveM’s Covid-19 Forecast Model Evaluations
  https://github.com/stevemcconnell/covid19-forecast-evaluations

• SteveM’s detailed articles on Medium
  https://stevemcc.medium.com/

• CDC forecast hub
  https://viz.covid19forecasthub.org/

Software Estimation Resources

• SteveM’s website (link to SteveM’s estimation book)
  https://www.stevemcconnell.com/

• Construx’s website—resources, online virtual classes, elearning classes
  https://www.construx.com/
Discussion