Comments on
“Data Models vs. Knowledge Models in Forecasting: The Way Forward”

J. Scott Armstrong

The Wharton School, U. of Pennsylvania, PA, jscott@upenn.edu

Kesten C. Green

University of South Australia Business School and Ehrenberg-Bass Institute, Adelaide, SA, Australia. kesten.green@unisa.edu.au

Presented at
The M4 Conference: *Advances in Forecasting*

December 10, New York

M4-Competition-slides-12-04-18.pptx

Procedure for this talk

- I often make mistakes, and I appreciate it when people inform me about them.
- This is the first time I’ve given this talk, so there may be mistakes.
- I will not be able to answer questions during the talk, but hopefully will be able near the end.
- Please write suggestions and give them to me at the end of the talk.

Success of the first M-competition

- Until the early 1980s, statisticians dominated forecasting.
  - They published complex untested methods.
- The first M-Competition was the first large-scale experimental test of the accuracy of alternative extrapolation methods
  - It found minor differences in accuracy and small improvements relative to the no-change model.
  - It also found gains from using simple methods and from combining forecasts from several validated methods.
  - Most importantly, it emphasized the need for *experimental tests of out-of-sample forecast accuracy*.

An outlet for experimental forecasting research

- The *Journal of Forecasting* was founded in 1982 to publish experimental research that tested the predictive validity of alternative methods.
- The M-Competition provided an ideal example.
- Thanks to the M-competition (and to papers by Nobel Prize winners), the 1982-83 citation impact factor for the *JoF* was 7th highest among business, management, and planning journals.

Data models vs. knowledge models

We define *data models* as those that reply on a statistical method to infer how the variable to be forecast is related to potential *predictor* variables.

Knowledge models, on the other hand, use prior knowledge (experimental or domain) to specify *causal* variables, the directions of effects, and effect sizes, and then use prior knowledge about forecasting methods to specify models.

Can you imagine anything that might convince you that data models may harm forecasting accuracy?

- Please spend one minute to write what would change your opinion.
- I am unable to change your beliefs. Only you can do that.
- I only change my opinions after reading about experimental research findings in scientific papers.

In conclusion

Do **not** use data models for forecasting.
Combining should only be used for validated methods.

Why not?

Let me count the whys.

Why? Reason 1: Data Models Violate the Golden Rule of Forecasting

"Be conservative by using prior knowledge about the situation and forecasting methods."

1. The Golden Rule paper provides 28 evidence-based guidelines.
2. A review of evidence identified 105 papers with 150 experimental comparisons; all supported the guidelines.
3. Ignoring a single guideline increased forecast error by more than 40% on average.

Data models ignore prior knowledge about forecasting methods and the situation.

Knowledge models

Benjamin Franklin proposed what we now call “knowledge models”.

1. Prior domain knowledge can be used to specify:
    - important causal variables,
    - the direction of the effects,
    - and the magnitude of the relationships.
2. There is no limit on the number of causal variables.
3. Equal weights are often more realistic than regression weights, *especially with many variables and where prior knowledge about effect sizes is poor.*
4. Knowledge models are expected to be more accurate than data models whenever experimental or domain knowledge about the situation is avaliable.

Combine Prior Knowledge with Extrapolation Methods

1. Extrapolation typically uses no causal variables, but domain knowledge about causality can be included for long-term extrapolations (i.e., modified based on experts' domain domain knowledge.)
2. Moreover, experimental knowledge exists about which forecasting methods work best under given conditions.
3. Rule-based forecasting uses domain knowledge to combine extrapolations based on 28 conditions of the data. It uses 99 simple rules to select the most accurate forecast for each of 18 features. For six-year ahead forecasts, the *ex ante* forecasts errors led to a 42% reduction vs. equal weights combining. **Rule-Based Forecasting: Development and Validation of an Expert Systems Approach to Combining Time Series Extrapolations**

Using Prior Knowledge with Extrapolation Methods: Examples

1. Decomposition of time-series by level and change, Tessier, *Journal of Business Research*, *68* (2015), 1755-1758.

2. Decomposition by Causal Forces: A Procedure for Forecasting Complex Time Series, Armstrong, Collopy & Yokum, *International Journal of Forecasting*, *21* (2005), 25-36

3. Causal Forces: Structuring Knowledge for Time-series Extrapolation, Armstrong & Collopy, *Journal of Forecasting, 12*, (1993), 103-115.

Why? Reason 2: Data models lack theoretical support

Einhorn (1972) refers to data models as "alchemy".
(**Alchemy in the behavioral sciences**. 1972)

We could find no theoretical support for data models as of 2018.

Why? Reason 3: Data models violate Occam's razor

1. A review of 32 studies found 97 comparisons between simple and complex methods. (Simple versus complex forecasting: The evidence.)
   a. None provided evidence that complexity improves forecast accuracy.
   b. On the contrary, complexity increased forecast error by an average of 27% in the 25 papers with quantitative comparisons.
2. All of the validated methods of forecasting are simple.
   (See the Methods checklist at forprin.com.)
3. Complexity causes minor damage for short-term extrapolation tests.
4. The harm increases for long-range forecasts, and when large changes in conditions are possible.

Why? Reason 4: Forecasting Software does not provide reliable findings

McCullough, B. D. (2000). Is it safe to assume that software is accurate? *International Journal of Forecasting*,16, 349-357.

His first paragraph is “No.”

Why? Reason 5: Data mining *enables* advocacy leading to unscientific (unethical?) practices

1. Advocates can use data models to provide forecasts that are favored by their clients… which helps researchers to get grants.
2. Researchers can easily obtain statistical significance to justify their models. (All findings become statistically significant if the sample sizes are large enough.)
3. Not surprisingly, then, statistical significance has been shown to harm science and decision-making.

Why? Reason 6: Data Models Violate the Guidelines for Regression analysis

We rated Data Models on the *Checklist for Using Regression Analysis*.

Data models violated at least 16 of the 18 guidelines.

No need to rely on our ratings. Do it yourself in about five minutes.

Data models should not be used for forecasting

1. Data models have been used by forecasters for over half a century.
2. Despite enormous expenditures, we could find no scientific evidence to support their use under any conditions
3. The traditional way of contributing to scientific knowledge is to test principles, such as *combining forecasts* within validated methods and then combining those across across different types of methods. This has led to error reductions of over 50%. (For evidence, see PollyVote.com.)
4. Said another way, scientific advances can only be made by "standing on the shoulders of giants."
5. We can only explain the continued fascination with data models by the *Rainmaker Theories* (from *Long-Range Forecasting*)

Rain Maker Theories

1. “The rainmaker gets so involved with the dance that he sometimes forgets that he has to make it rain.”

2. “Yes, I know it didn’t rain—but didn’t you like the dance?”
3. The successful rainmaker is the one who can convince his client that he didn’t want rain—he wanted to watch the dance.

To know the way forward, you must know where you started

Use prior knowledge about validated methods and principles of forecasting.
(see **Forecasting Methods and Principles: Evidence-Based Checklists**.)

The most powerful forecasting generalization to date, is to combine forecasts *within methods*, then to combine across these combined forecasts *across different validated methods*.

For continuing developments in forecasting, see ForecastingPrinciples.com (forprin.com) To date, there have been over 15 million visits to this site.

These slides will be available at ForPrin.com and ResearchGate.

Write action steps for yourself and suggestions for us

Action steps On paper: promises for yourself)

Suggestions (on paper please)

Questions?

You can also contact us by email: jscott@upenn.edu or kesten.green@unisa.edu.au