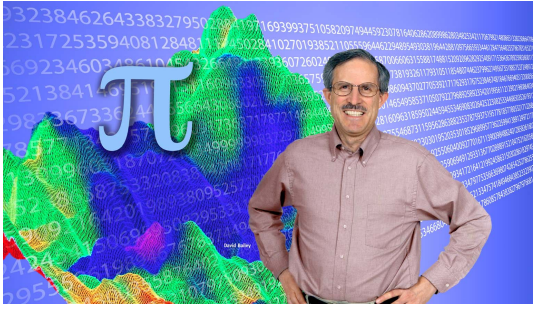
High-performance computing (and the dangers of overfitting) in finance

David H. Bailey
http://www.davidhbailey.com
Lawrence Berkeley National Lab (retired) and University of California, Davis



Artificial intelligence and big data in finance

According to a Bloomberg report, Wall Street is entering a new era:

> *The fraternity of bond jockeys, derivatives mavens and stock pickers who've long personified the industry are giving way to algorithms, and soon, artificial intelligence.*
>
> *Banks and investment funds have been tinkering for years, prompting anxiety for employees. Now, firms are rolling out machine-learning software to suggest bets, set prices and craft hedges. The tools will relieve staff of routine tasks and offer an edge to those who stay. But one day, machines may not need much help.*

The message to the finance world is clear: Prepare to adapt, or face irrelevance.

- S. Kishan, H. Son and M. Rojanasakul, "Robots are coming for these Wall Street jobs," https://www.bloomberg.com/graphics/2017-wall-street-robots/.

Challenging times for hedge funds

Many hedge funds are facing withdrawals and closures, as returns have lagged the market.

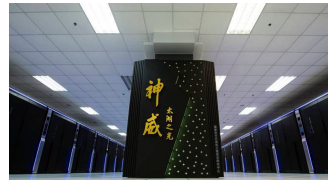
Yet some hedge funds do consistently beat the market! According to *Institutional Investor's Alpha*:

> *Five of the six largest firms in this year's ranking rely all or mostly on computers and algorithms to make their investment decisions, a theme that has increasingly played a role in the top-100 ranking over the past few years. And all five posted asset increases last year.*

- "Which hedge funds actually beat the market?," Mathematical Investor, 11 Aug 2017, http://mathinvestor.org/which-hedge-funds-actually-beat-the-market.

High-performance computing in science and technology

A large community has been using highly parallel systems and graphics processing unit (GPU) accelerators in a broad range scientific and engineering disciplines:



- Weather forecasting.
- Climate modeling.
- Environmental data processing and analysis.
- Automobile engineering.
- Aerospace engineering.
- Astrophysics and cosmology.
- Biology.
- Pharmaceutical research.
- Consumer product development.

The annual Supercomputing Conference typically draws 12,000 or more attendees and features exhibits with hundreds of millions of dollars of state-of-the-art systems.

High-performance computing in finance

Highly parallel and GPU-accelerated computer systems are rapidly becoming an essential staple of high-end finance:

- Performing large-scale factor and correlation analyses on huge financial datasets.
- Portfolio optimization — optimize thousands or millions of individual portfolios in real time, based on trading results.
- Valuation of derivatives — dynamically evaluate hedge strategies in real time.
- Simulation of portfolio behavior over a broad range of future scenarios.
- Large-scale alternative data analysis (satellite photos, etc.).
- Machine learning applications using sophisticated algorithms and large datasets.

DANGER AHEAD

Artificial intelligence, big data and high-performance computing can generate nonsense faster than ever before!

Details to consider:

- Are the theoretical models realistic?
- Is the computer implementation debugged?
- Are the statistical methods and data analysis techniques appropriate?
- Is there sufficient data to draw statistically reliable conclusions?
- Can other researchers reproduce the findings?

- V. Stodden, M. McNutt, D. H. Bailey, E. Deelman, Y. Gil, B. Hanson, M. A. Heroux, J. P. A. Ioannidis and M. Taufer, "Enhancing reproducibility for computational methods," *Science*, vol. 356, no. 3617 (3 Dec 2016), 1240–1241.

Reproducibility crises in biomedicine, psychology and economics

- In 2011, Bayer researchers reported that they were able to reproduce only 17 of 67 pharma studies.
- In 2012, Amgen researchers reported that they were able to reproduce only 6 of 53 cancer studies.
- In August 2015, the Reproduciblity Project in Virginia reported that they were able to reproduce only 39 of 100 psychology studies.
- In September 2015, the U.S. Federal Reserve was able to reproduce only 29 of 67 economics studies.

Reproducibility Project staff

Credit: NY Times

Email exchange between DHB and a finance colleague

Email from DHB to finance colleague, 10 June 2013:

One thing that has always puzzled me about the financial world is the following sort of thing: [press examples cited]. Excuse me for being "dumb," but this sort of thing seems to me to be outright nonsense. ... When people like those above say that they "know" where the stock market is heading, this cannot have any scientific basis. ...

So why doesn't somebody blow this whistle on this sort of thing? Am I missing something?

Response from finance colleague to DHB, 17 June 2013:

It is not a dumb question at all. It is a question I have struggled with and which answer makes me an unhappy man. The truth is, most people in this industry are charlatans. They do not have any particular model or theory to understand the world. They are not scientists. ...

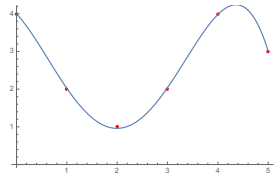
I completely agree with your assessment. The amount of nonsense ... is incredible.

What is backtest overfitting?

Backtest: Testing a model based on historical market data.
Backtest overfitting: Statistical overfitting of historical market data.

Examples:

- Proposing a theoretical model that inherently possesses a higher level of complexity than the backtest data.
- Using a computer to try millions or billions of variations of a model or strategy on historical market data, and then only presenting results from the variation that works best.
- Constructing an investment fund by using a computer to explore millions or billions of weighting factors, then only marketing the one that works best.

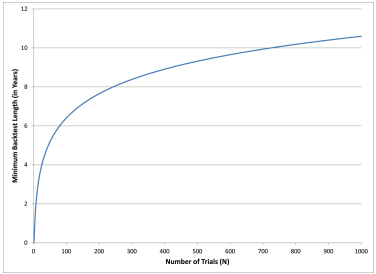


Fitting six data points (almost perfectly!) with a fourth-degree function.

When a computer can analyze many variations of a fund or strategy on a fixed dataset, it is almost certain that the "optimal" selection will be statistically overfit.

How easy is it to overfit a backtest? Very!

- If only 2 years of daily market data are available, then no more than 7 variations should be tried.
- If only 5 years of daily market data are available, then no more than 45 variations should be tried.



Backtest overfitting is now thought to be the principal reason why financial funds and strategies that look great on paper often disappoint in practice.

- D. H. Bailey, J. M. Borwein, M. Lopez de Prado and J. Zhu, "Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance," *Notices of the American Mathematical Society*, May 2014, 458–471.

Letters to clients: An absurd investment scheme

- A financial advisor sends letters to $5,120 = 5 \times 2^{10}$ prospective clients, with 2560 predicting a certain stock will go up, and 2560 predicting it will go down.
- One month later, the advisor sends letters only to the 2560 investors who were previously sent the correct prediction, with 1280 letters predicting a certain stock will go up, and 1280 predicting it will go down.
- After ten months, the final five investors will have been sent ten consecutive spot-on predictions!

This strategy is absurd, even fraudulent, because the final five investors are not told of the 10,235 other letters with different predictions.

But why is promoting a statistically overfit fund or strategy, where investors are not informed of the millions of failed computer trials, any different?

Other potential abuses of computing in finance

Many in the finance world remain disappointingly silent with regards to those in the community who, knowingly or not:

- Engage in "p-hacking" to inflate the statistical significance of their results.
- Fail to disclose the number of models or variations that were tried in their study.
- Misuse probability theory and statistical methods.
- Present misleading charts and graphs.
- Use questionable pseudomathematical jargon.
- Make vague claims that do not permit rigorous testing and falsification.
- Withhold key details, so other researchers cannot test or reproduce their results.

As we wrote in a recent paper:
Our silence is consent, making us accomplices in these abuses.

Supercomputers and GPUs can't save us

Lessons from the high-performance computing world:

- Extra speed may just produce nonsense faster.
- Extra speed may cover up the usage of inaccurate and unrealistic models.
- Extra speed may cover up the usage of inefficient, ineffective, inappropriate or numerically unstable algorithms.
- Extra speed may tempt analysts into "optimizing" over many model variations, which almost certainly will result in statistical overfitting.
- Extra speed may mislead management into thinking that new hardware is more important than knowledgeable, experienced human analysts.

Thanks!

This talk is available at:
http://www.davidhbailey.com/dhbtalks/dhb-aidata-2017.pdf