Forecasting Conflict
Lecture 1
Technical Political Forecasting: An Overview

Philip A. Schrodt

Parus Analytical Systems
schrodt735@gmail.com

Graduate School of Decision Sciences
University of Konstanz
14 - 17 October 2013

Overview of the series of lectures

1. Prediction: contemporary opportunities and challenges
2. Conflict forecasting in U.S. government projects
3. Event data and GDELT
4. Forecasting: Conventional time series approaches
5. Forecasting: Sequence and clustering approaches

Overview for First Lecture

- Justification of prediction
- Seven opportunities
- Seven challenges

Approach of the lectures

- Breadth, not depth!—this is more of a "bird's eye view"
  - (But not—I repeat, **not!**—a "god's eye view"!)
- A guide to vocabulary[ies], approaches and what you need to know
  - you can then follow up on all of this material in detail. If I can look it up, you can look it up
- Emphasis on practical applications: Some of the slides are recycled from presentations I've given in the U.S. policy community
  - This is a feature, not a bug
- This is the departure lounge, not the baggage claim
- All of the slides are available at

  `http://eventdata.parusanalytics.com/presentations.html`

The Debate

▸ ARGUMENT PRINT | TEXT SIZE | EMAIL | SINGLE PAGE

Why the World Can't Have a Nate Silver

The quants are riding high after Team Data crushed Team Gut in the U.S. election forecasts. But predicting the Electoral College vote is child's play next to some of these hard targets.

BY JAY ULFELDER | NOVEMBER 8, 2012

Vs.

▸ ARGUMENT PRINT | TEXT SIZE | EMAIL | SINGLE PAGE

Predicting the Future Is Easier Than It Looks

Nate Silver was just the beginning. Some of the same statistical techniques used by America's forecaster-in-chief are about to revolutionize world politics.

BY MICHAEL D. WARD , NILS METTERNICH | NOVEMBER 16, 2012

Whatever is not forbidden is mandatory

The old challenge: “Prediction is not scientific”

Whatever is not forbidden is mandatory

The old challenge: “Prediction is not scientific”
Huh????

Whatever is not forbidden is mandatory

The old challenge: “Prediction is not scientific”
Huh????

The new challenge: “Political science (and economics) is not scientific because it can’t predict”

- See various op-eds in *New York Times* and related venues over the past eighteen months. Despite Nate Silver’s predictions.
- See complete suspension of US NSF Political Science program
- +<4> The sting of which has been somewhat mollified by subsequent efforts by the same people to suspend the entire US government

Problems with qualitative approaches

Tetlock: Experts typically do about as well as a “dart-throwing chimp”

Problems with qualitative approaches

Tetlock: Experts typically do about as well as a "dart-throwing chimp"

Except for television pundits, who do even worse. Ask President Romney.
The media want things to be dramatic. "We're all going to die! Details follow *American Idol*"

Problems with qualitative approaches

Tetlock: Experts typically do about as well as a "dart-throwing chimp"

Except for television pundits, who do even worse. Ask President Romney.
The media want things to be dramatic. "We're all going to die! Details follow *American Idol*"

Qualitative *theory* isn't much better:
Remember the hegemonic US seizure of undefended Canadian and Mexican oil fields in response to the 1973 OPEC oil embargo?

SMEs and the “narrative fallacy”

SME = “subject matter expert”

Hegel: the owl of Minerva flies only at dusk

Taleb (*Black Swan*): seeking out narratives is an almost unavoidable cognitive function and it generates a dopamine hit

This is your brain on narratives

Problems with quantitative approaches

Ward, Greenhill and Bakke (2010): Models based on significance tests don't predict well because that is not what a significance test is supposed to do.

Gill, Jeff. 1999. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly* 52:3, 647-674.

Problems with quantitative approaches

Ward, Greenhill and Bakke (2010): Models based on significance tests don't predict well because that is not what a significance test is supposed to do.

Gill, Jeff. 1999. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly* 52:3, 647-674.

The norm in political science has been to do full-sample evaluation, whereas the norm in machine-learning has been split-sample, which is usually more robust and is certainly more credible

Role of prediction for logical positivists

Hemple: "Explanation" in the absence of prediction is "prescientific"

- Critical case: astrology vs astronomy
- More generally, mythological accounts provide "explanation" [Quine]

Prediction was simply assumed to be a defining characteristic of a good theory until relatively recently

Arguably, no philosopher of science prior to the mid-20th century would find the frequentist-based "explanation" emphasized in contemporary political science even remotely justified

- Leaving aside that frequentism is logically inconsistent and has been characterized in Meehl (1978) as "a terrible mistake, basically unsound, poor scientific strategy, and one of the worst things that ever happened in the history of psychology"
- Hey, dude, tell us what you *really* think...
- But that is another lecture...

Explanation, continued

Philosophers of science have long suspected that it is possible to have a seemingly sound explanation of a phenomenon that confers no predictive leverage over the phenomenon (Nagel, 1961; Toulmin, 1961). For instance, plate tectonics theory is the received explanation for earthquakes, but it confers no ability to generate accurate predictions about when earthquakes will occur. Conversely, it is possible to have remarkable predictive accuracy that rests on a deeply flawed framework. Ancient astronomers generated predictively powerful celestial charts even though they didn't have the faintest idea what planets or stars were.

. . .

How patient should we be with low-predictive-accuracy theories? When should we tune out the theorists and go with algorithms that no more understand world politics than ancient astronomers understood celestial motion? We have no off-the-shelf answer, but we resonate to Lakatos's (1970) rule for distinguishing degenerative from progressive research programs: forgive patch-up operations only if they inspire testable propositions that pan out. [Tetlock, 2013]

Additional issues in explanation vs. theory

Hume: the problem of induction

- Farmer's cat vs. farmer's turkey

Friedman: unreasonable assumptions are justified provided the predictions are accurate

- Justification for rational choice models
- Issue: the "provided predictions are accurate" part tends to be forgotten, and is far too often replaced with "provided I think the assumptions are elegant and/or make my life easier"

Success without theory: Gothic cathedrals

Note that these issues affect *observational* studies but not *experimental* studies, which is why experiments are used whenever possible.

Kahneman et al: people are really bad at statistical reasoning

- Everyone, including statisticians unless they focus very hard
- Example: managed mutual funds, which both theory and evidence indicate cannot work
- Example: opposition to "evidence based medicine" in the US, with a preference for clinical intuition even when this has been demonstrated to be less effective
- Probabilitistic weather forecasts seem to be the one major exception: rain likelihood, hurricane tracks

The Necessity of Prediction in Policy

Feedforward: policy choices must be made in the present for outcomes which may not occur for many years

Planning Times: even responses to current conditions may require lead times of weeks or months

[More on this tomorrow]

“Schrodt should do everything in ‘sevens’”

Opportunities

- Totalitarian law of the universe: whatever is not forbidden is mandatory. Prediction *is* scientific [now]
- Data: small, big, fast
- You can't solve everything with more machine cycles, but it ~~never~~ rarely hurts
- Successful large-scale projects: PITF, ICEWS, ACE, ENCoRe
- (Mostly) Convergent models
- Location, location, location
- Open source, open access, open collaboration

Challenges

- Determining credible metrics
- Black swans
- Heterogeneous environments
- Absence of theories indicating what is and is not predictable
- Pournelle's Law: no task is so virtuous that it will not attract idiots
- Ethical concerns

Multi-disciplinary challenges

Big Data: Machine learning researchers routinely use social science data to construct models. Many of these achieve high accuracy in split-sample tests, to the point where these researchers simply assume that things are predictable.

IARPA ACE "Good Judgment Project" (Tetlock): While *most* forecasters do no better than chance, a small number of "super forecasters" perform significantly over long periods of time and large numbers of questions. Furthermore these individuals exhibit common characteristics and strategies, and to a limited extent, these can be taught. Forthcoming article in *Economist: The Year 2014*

Early work (1980s) in "expert systems" for classification problems showed the systems tended to be about 10% more accurate than humans and achieved this accuracy using less information.

The Forecaster's Quartet

- Nassem Nicholas Taleb. *The Black Swan*
  (most ~~entertaining~~ obnoxious)
- Daniel Kahneman. *Thinking Fast and Slow*
  (30 years of research which won Nobel Prize)
- Philip Tetlock. *Expert Political Judgment*
  (most directly relevant)
- Nate Silver. *The Signal and the Noise*
  (high level of credibility after perfect 2012 electoral vote predictions)

Data!


Science
11 February 2011 | $10
scientific
data
information
climate
research
science
new
analysis
visualization
researchers
access
many
human
knowledge
example
AAAS


PRIO


ACLED
Armed Conflict Location & Event Dataset


Minorities at Risk

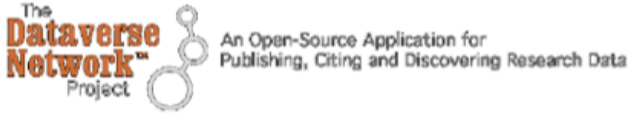

CORRELATES OF
WAR


UPPSALA
UNIVERSITET


PENNSTATE
1855


UCDP
GEOREFERENCED EVENT DATASET


CIRI Human Rights Data Project
www.humanrightsdata.org


The
Dataverse
Network
Project
An Open-Source Application for
Publishing, Citing and Discovering Research Data


Polity
IV


GTD
Global Terrorism Database


Universität
Konstanz

Though this may be going a little far...

WIRED MAGAZINE: 16.07

SCIENCE : DISCOVERIES

The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

By Chris Anderson 06.23.08

Illustration: Marian Bantjes

Computing power

Computing Power

Control Data Corporation 3600
(ca.1965)
32 K (48-bit) RAM memory
1 processor
~1-million operations per second
Output: line printer

Computing Power

Control Data Corporation 3600
(ca.1965)
32 K (48-bit) RAM memory
1 processor
~1-million operations per second
Output: line printer

Penn State High Performance Computing Facility
15 cluster computers
100 to 2000 2.66 Ghz processors in each cluster
~50 Gb RAM accessible to each processor
130 Tb disk space
4 interactive visualization rooms

Computing Power

Control Data Corporation 3600
(ca.1965)
32 K (48-bit) RAM memory
1 processor
~1-million operations per second
Output: line printer

Penn State High Performance Computing Facility
15 cluster computers
100 to 2000 2.66 Ghz processors in each cluster
~50 Gb RAM accessible to each processor
130 Tb disk space
4 interactive visualization rooms

Motorola Razr
16 Gb RAM memory
Dual processor
~500-million operations per sec
540 x 860 color display

Open Source Software

MySQL

python

mozilla Firefox

ubuntu

source forge

T R

perl

Ruby PROGRAMMING Language

WordPress.com

APACHE HTTP SERVER

R

LibreOffice The Document Foundation

GNU run free run GNU

stackoverflow

FREE SOFTWARE FOUNDATION

Computationally-intensive methods

- Bayesian estimation using Markov chain Monte Carlo methods
- Bayesian model averaging ("*AJPS*-as-algorithm")
- random forest models
- large-scale textual databases
- machine translation
- geospatial visualization
- real-time automated coding
- remote sensing data such as nightlight density

Large Scale Conflict Forecasting Projects

- State Failures Project 1994-2001
- Joint Warfare Analysis Center 1997
- FEWER [Davies and Gurr 1998]
- Center for Army Analysis 2002-2005
- Swiss Peace Foundation FAST 2000-2008
- Political Instability Task Force 2002-present
- DARPA ICEWS 2007-present
- IARPA ACE and OSI
- Peace Research Center Oslo (PRIO) and Uppsala University UCDP models

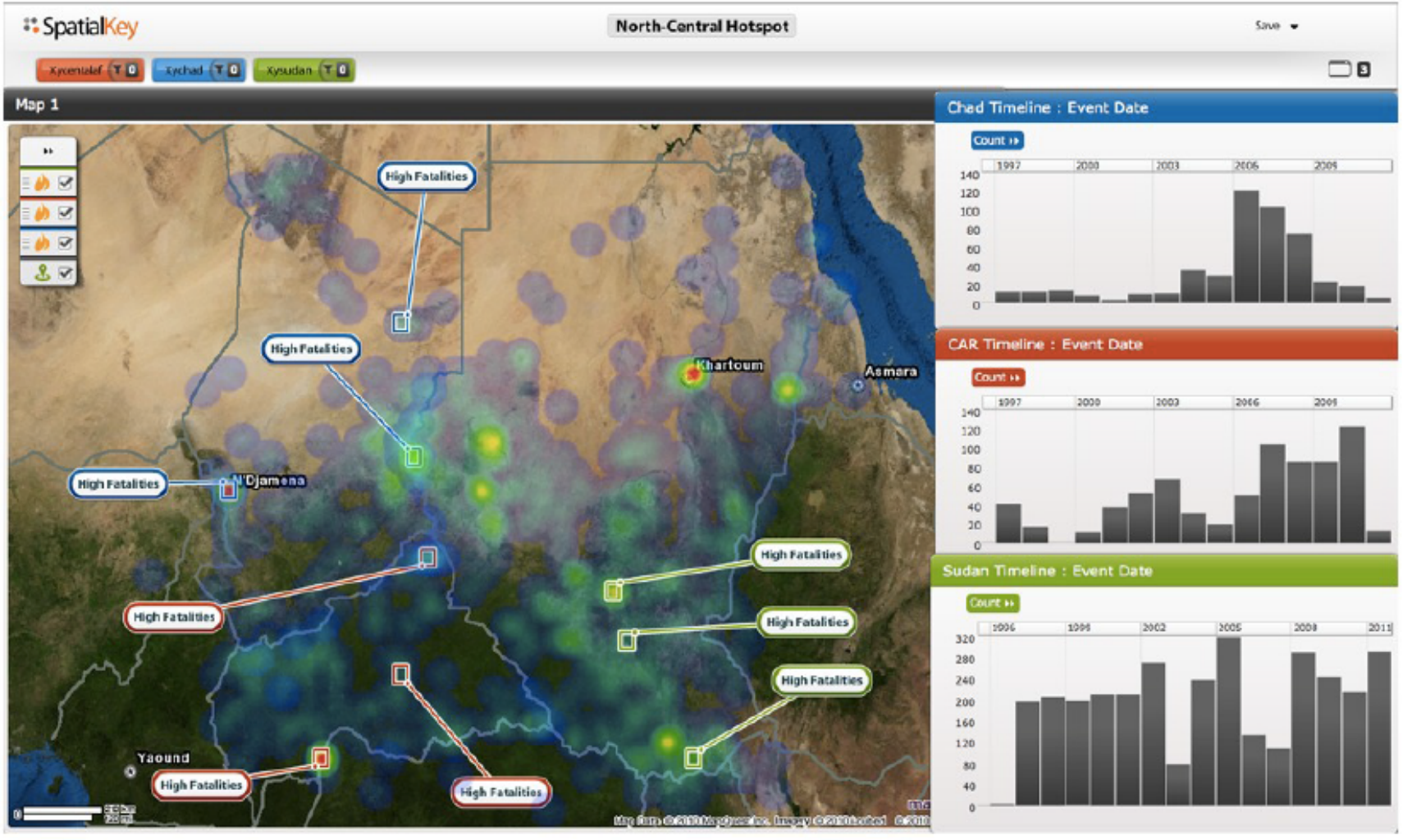
(much more on this tomorrow)

Convergent Results

- Most models require only a [very] small number of variables
- Indirect indicators—famously, infant mortality rate as an indicator of development—are very useful
- Temporal autoregressive effects are huge: the challenge is predicting onsets and cessations, not continuations
- Spatial autoregressive effects—"bad neighborhoods"—are also huge
- Multiple modeling approaches generally converge to similar accuracy
- 80% accuracy—in the sense of AUC around 0.8— in the 6 to 24 month forecasting window occurs with remarkable consistency: few if any replicable models exceed this, and models below that level can usually be improved
- Measurement error on many of the dependent variables—for example casualties, coup attempts—is still very large
- Forecast accuracy does not decline very rapidly with increased forecast windows, suggesting long term structural factors rather than short-term "triggers" are dominant. Trigger models more generally do poorly except as *post hoc* "explanations."

Location, location, location!

ACLED Geospatial



UCDP Geospatial

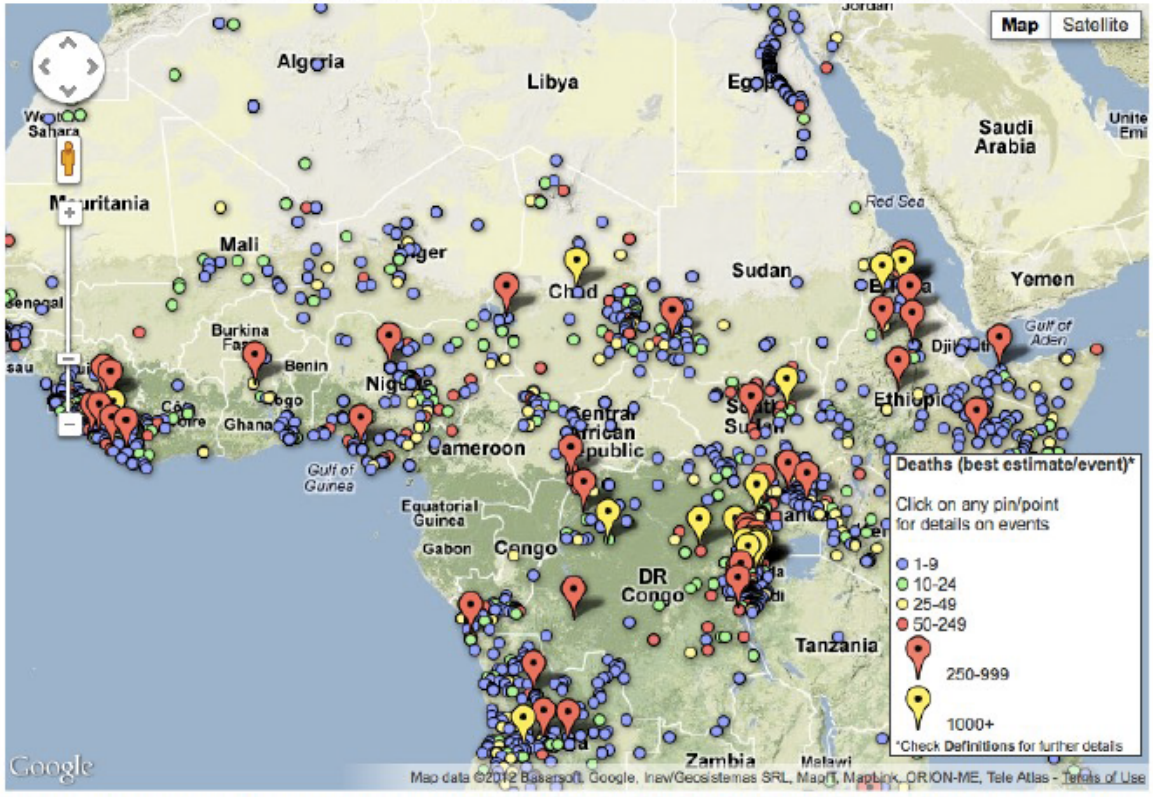
Welcome to the UCDP GED - Uppsala Conflict Data Program's Georeferenced Event Dataset

The interactive map below covers all the locations geocoded as part of GED. ***Drag and move the red and grey thumbs of the slider below to filter for particular intervals, years, or even particular months or days.*** *Use the filters to select only one of the three UCDP types of violence. Clicking on each dot or pin will display information regarding events taking place in those respective locations. Download data here.*

1989 2010

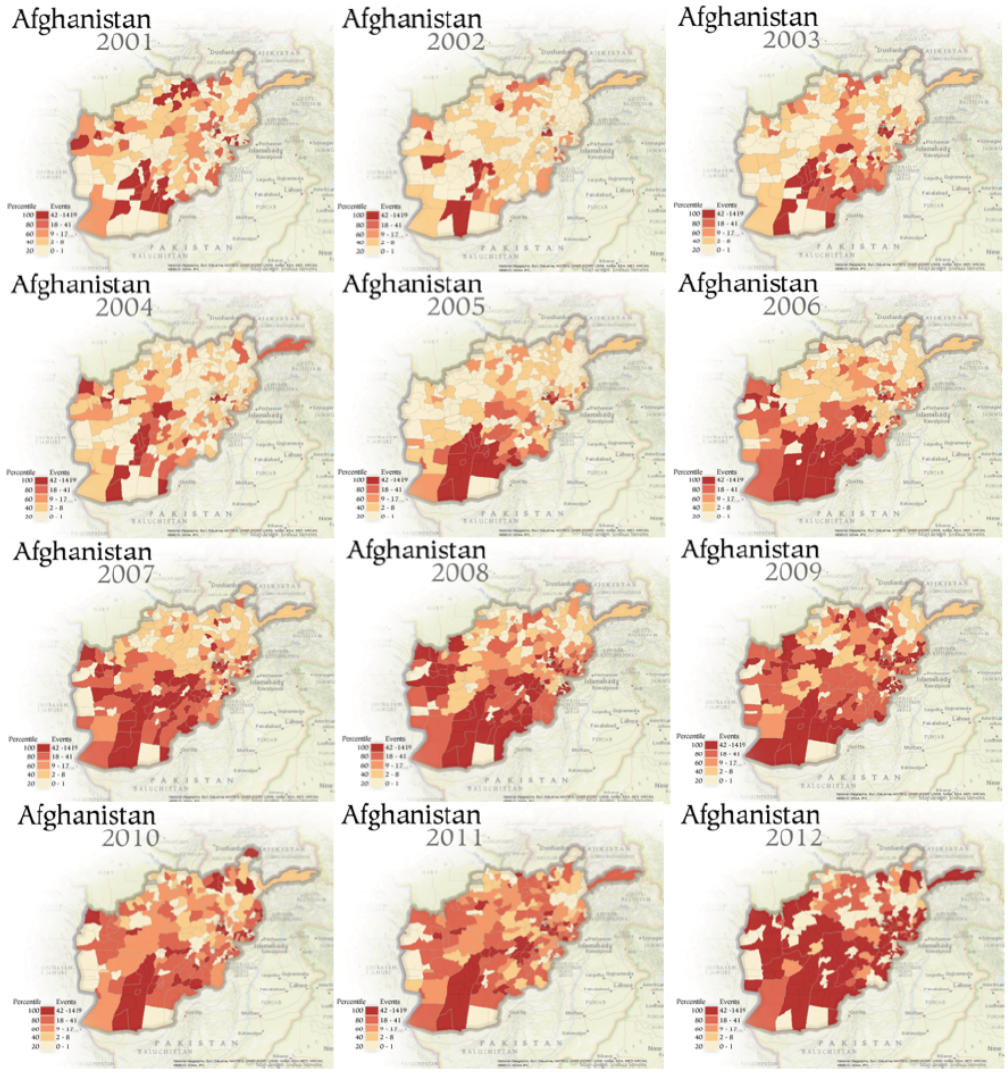
Filter for:

All violence | State-based conflicts only | Non-state conflicts only | One-sided conflicts only



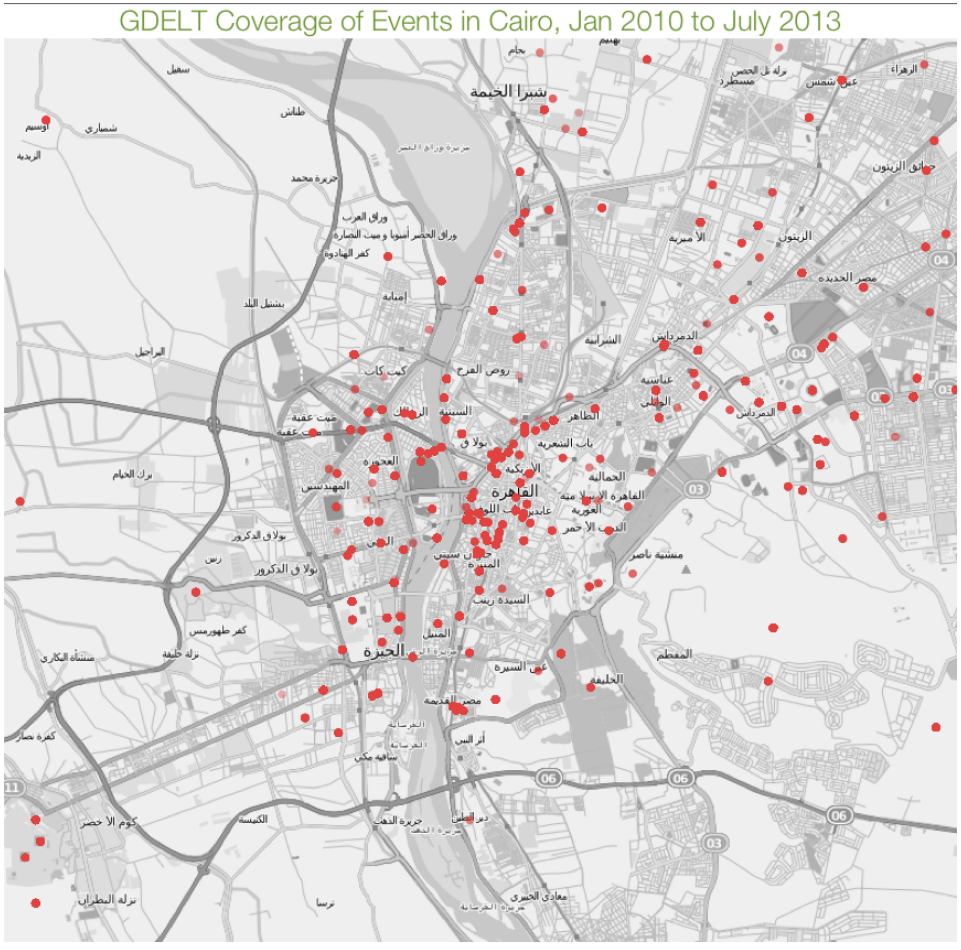
The GED is the product of two and a half years of work at the **Department of Peace and Conflict Research, Uppsala University**. The UCDP GED contains conflict data disaggregated spatially and temporally down to the level of the individual incidents of violence. For more details please see the About UCDP link above.

GDELT: Afghanistan, District-level Violence



Source: Jay Yonamine and Joshua Stevens, Penn State

GDELT: Cairo protests



Source: David Masad and Andrew Halterman of Caerus Analytics.

Open source, open access, open collaboration

- There is a strong if incomplete norm towards open sharing of data and methods
  - Unintended consequence: PITF "forecasting tournament" cannot be published in a major journal because it used proprietary data—the baseline data has 2,700 variables—that cannot be archived in replication sets. The results are, however, still available on SSRN.
  - The inability to share source texts is clearly a concern in news-report-based datasets such as GDELT and MID.
- By all available evidence, US government forecasting projects are using similar methodologies to those available in open sources; in fact they are probably lagging somewhat behind this
- We now have significant NGO and academic work, and an international "epistemic community" has developed around the topic.

CHALLENGES

Metrics

What is being predicted

- Probability of binary outcomes by a fixed date
- Quintile rankings of risk / probability-based “watch lists”
- Survival and hazard models
- Switching and phase models
- Networking—both social and geospatial—models

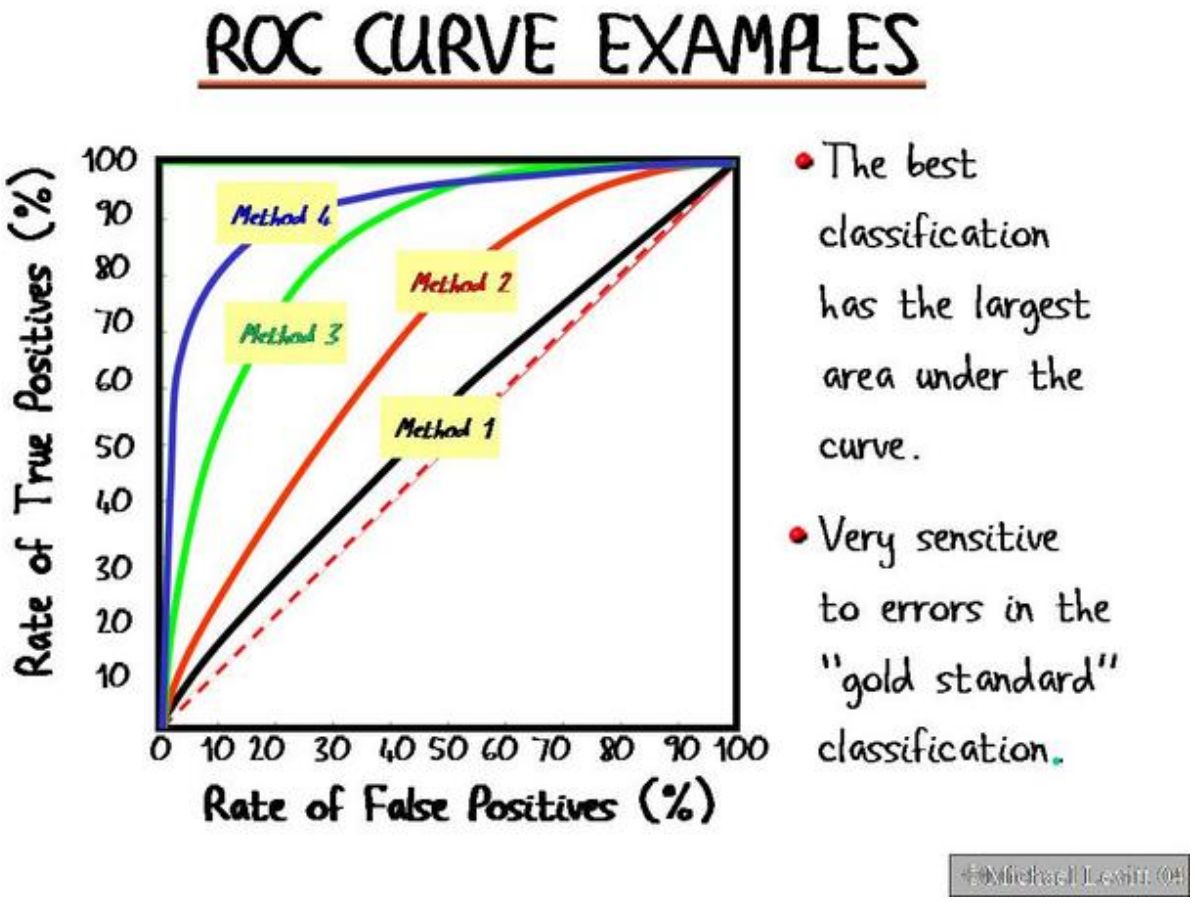
All of these can be used as input to ensemble methods

Classification Matrix

Relationships among terms

		Condition (as determined by "Gold standard")		
		Condition Positive	Condition Negative	
Test Outcome	Test Outcome Positive	**True Positive**	**False Positive** (Type I error)	Positive predictive value = $\frac{\Sigma \text{ True Positive}}{\Sigma \text{ Test Outcome Positive}}$
	Test Outcome Negative	**False Negative** (Type II error)	**True Negative**	Negative predictive value = $\frac{\Sigma \text{ True Negative}}{\Sigma \text{ Test Outcome Negative}}$
		Sensitivity = $\frac{\Sigma \text{ True Positive}}{\Sigma \text{ Condition Positive}}$	**Specificity** = $\frac{\Sigma \text{ True Negative}}{\Sigma \text{ Condition Negative}}$	

ROC Curve



Source: http://csb.stanford.edu/class/public/lectures/lec4/Lecture6/Data_Visualization/images/Roc_Curve_Examples.jpg

Separation plots

TABLE 4 Rearrangement (and Coloring) of the Data Presented in Table 1 for Use in the Separation Plot

Country	Fitted Value ($\hat{p}$)	Actual Outcome (y)
B	0.364	0
F	0.422	1
D	0.728	0
A	0.774	0
E	0.961	1
C	0.997	1

FIGURE 2 Separation Plot Representing the Data Presented in Table 1

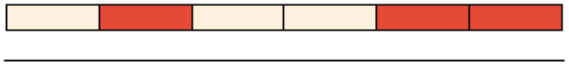
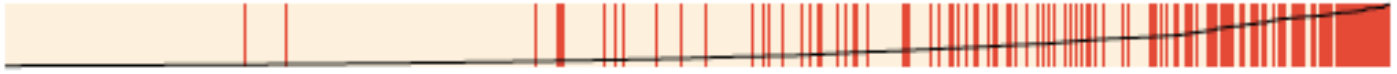


FIGURE 3 Separation Plot for a Larger Data Set

FIGURE 4 Adding a Graph of $\hat{p}$ to the Separation Plot



The Separation Plot: A New Visual Method for Evaluating the Fit of Binary Models

Brian Greenhill[1], Michael D. Ward[2], Audrey Sacks[3]

Article first published online: 20 JUN 2011

DOI: 10.1111/j.1540-5907.2011.00525.x

Issue

American Journal of Political Science

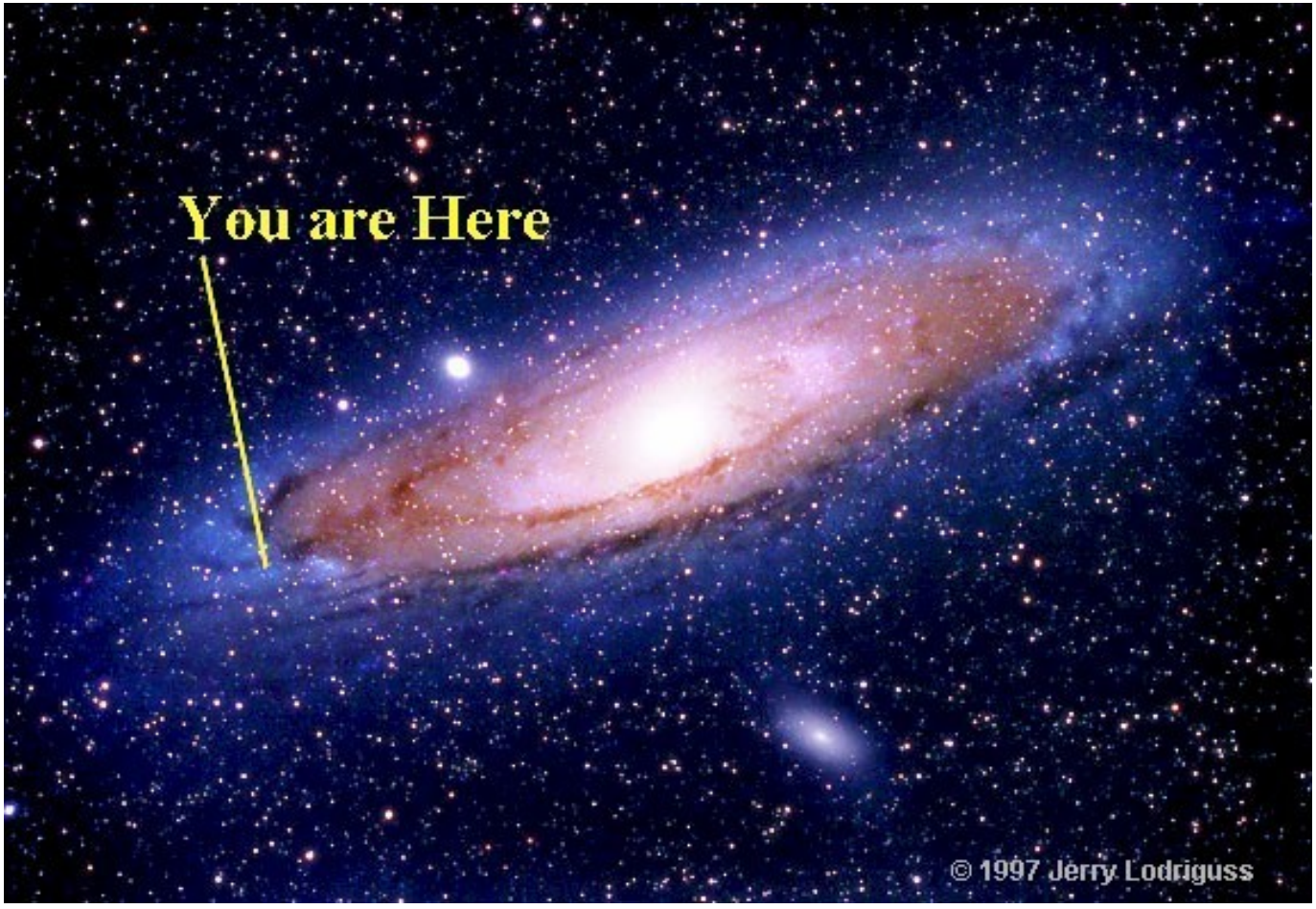
Volume 55, Issue 4, pages 991–1002, October 2011

And wait, there's still more!

- Recall / True Positive Rate/Sensitivity
- Precision / Positive predictive value (PPV)
- Specificity / True Negative Rate
- F1 score: harmonic mean of precision and recall
- Beier scores
- Posterior probabilities
- Proportional reduction of error or entropy

Black swans

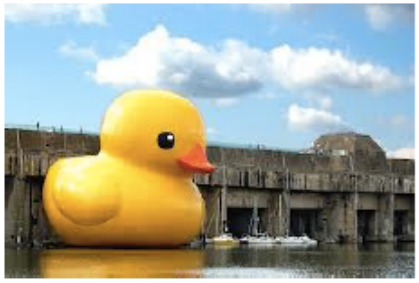
Ideal forecasting targets are neither too common nor too frequent



Good Judgment Project: look for events with a 10% probability

The Forecasting Zoo

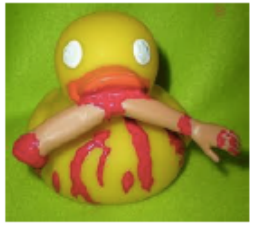
Ducks can be interesting...



Size

Variety

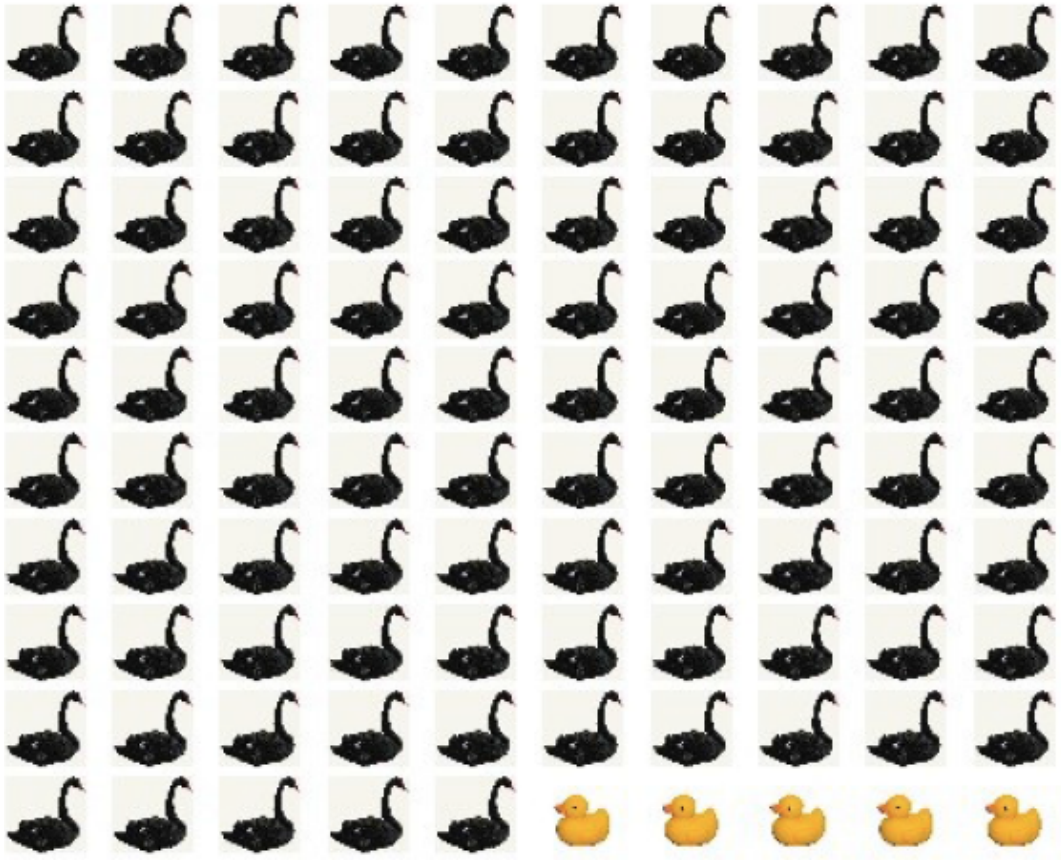
Quantity



Suspicious behaviors

And this is going too far…

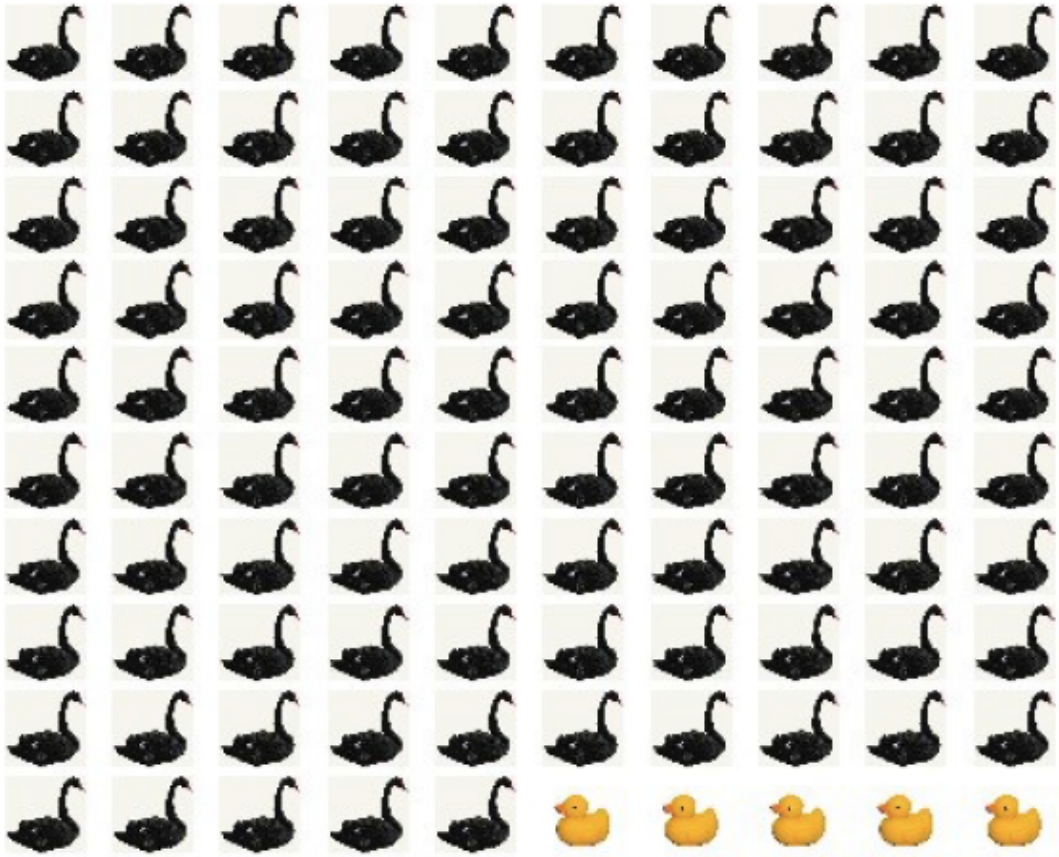
DARPA-World!



By definition, most black swans *will not occur*! So there is little point in investing a large amount of effort trying to predict them.

And this is going too far…

DARPA-World!



By definition, most black swans *will not occur*! So there is little point in investing a large amount of effort trying to predict them.
"Can your model predict a chemical attack by self-recruited Mexican jihadis working as rodeo clowns in Evanston, Wyoming? Why not?!"

Challenge: distinguishing black swans from rare events

Black swan: an event that has a low probability even conditional on other variables

Rare event: an event that occurs infrequently, but conditional on an appropriate set of variables, does not have a low probability

Medical analogy: certain rare forms of cancer appear to be highly correlated with specific rare genetic mutations. Conditioned on those mutations, they are not black swans.

Another important category: high probability events which are ignored. The "sub-prime mortgage crisis" was the result of the failure of a large number of mortgage which models had completely accurately identified as "sub-prime" and thus likely to fail. This was not a low probability event.
Upton Sinclair: It is hard to persuade someone to believe something when he can make a great deal of money not believing it.

Heterogeneous environments

- Per Pinker, Goldstein, Mueller, etc, is the system changing significantly while we are trying to model it? How far back are data still relevant?
- How different are various types of militarized non-state actors? For example, how much do al-Qaeda and international narcotics networks have in common?
- We are also using a more heterogenous set of forecasting methods, and probably do not understand their weak points as well as we understand those of regression-based models.
- Threats tend to occur in small number of "hot-spots"
  - Europe 1910-1945
  - Middle East 1965-present
  - Balkans in 1990s
  - Internal conflict in India

Note that all of these are complicated by rare events—some of which may be black swans—since it limits the number of observations we have on the dependent variable.

Changing nature of conflict

Threats in 1910

- "Gunboat diplomacy" was an accepted norm, as were elements of bellicose and social Darwinism
- Some competition occurred between approximate equals
- Mediation was *ad hoc* with no established international institutions
- Territorial change was credible

Threats in 2010

- Highly asymmetric distribution of military power
- Threats get almost immediate attention from potential mediators, including the UN
- Non-military sanctions are credible (Iraq, Iran)
- Territorial changes are rare and highly problematic

Will changes in the technological environment—internet, UAVs, various monitoring technologies—change probabilities?

Theory: what can and cannot be predicted?

Is astronomy scientific?

Astronomy generally has a very good record of prediction, and from the earliest days of astronomy, successful prediction has been a key legitimating factor

- relation between star positions and events like the Nile flood
- eclipses
- orbits
- Halley's comet
- precision steering of space-craft

Is astronomy scientific?

Astronomy generally has a very good record of prediction, and from the earliest days of astronomy, successful prediction has been a key legitimating factor

- relation between star positions and events like the Nile flood
- eclipses
- orbits
- Halley's comet
- precision steering of space-craft

Nonetheless, astronomy cannot predict, nor does it attempt to predict:

- solar flares, despite their potentially huge economic consequences
- previously unseen comets
- next nearby supernova

Determinism: The Pioneer spacecraft anomaly

"[Following 30 years of observations] When all known forces acting on the spacecraft are taken into consideration, a very small but unexplained force remains. It appears to cause a constant sunward acceleration of $(8.74 \pm 1.33) \times 10^{-10} m/s^2$ for both spacecraft."

Source: Wikipedia

Irreducible sources of error-1

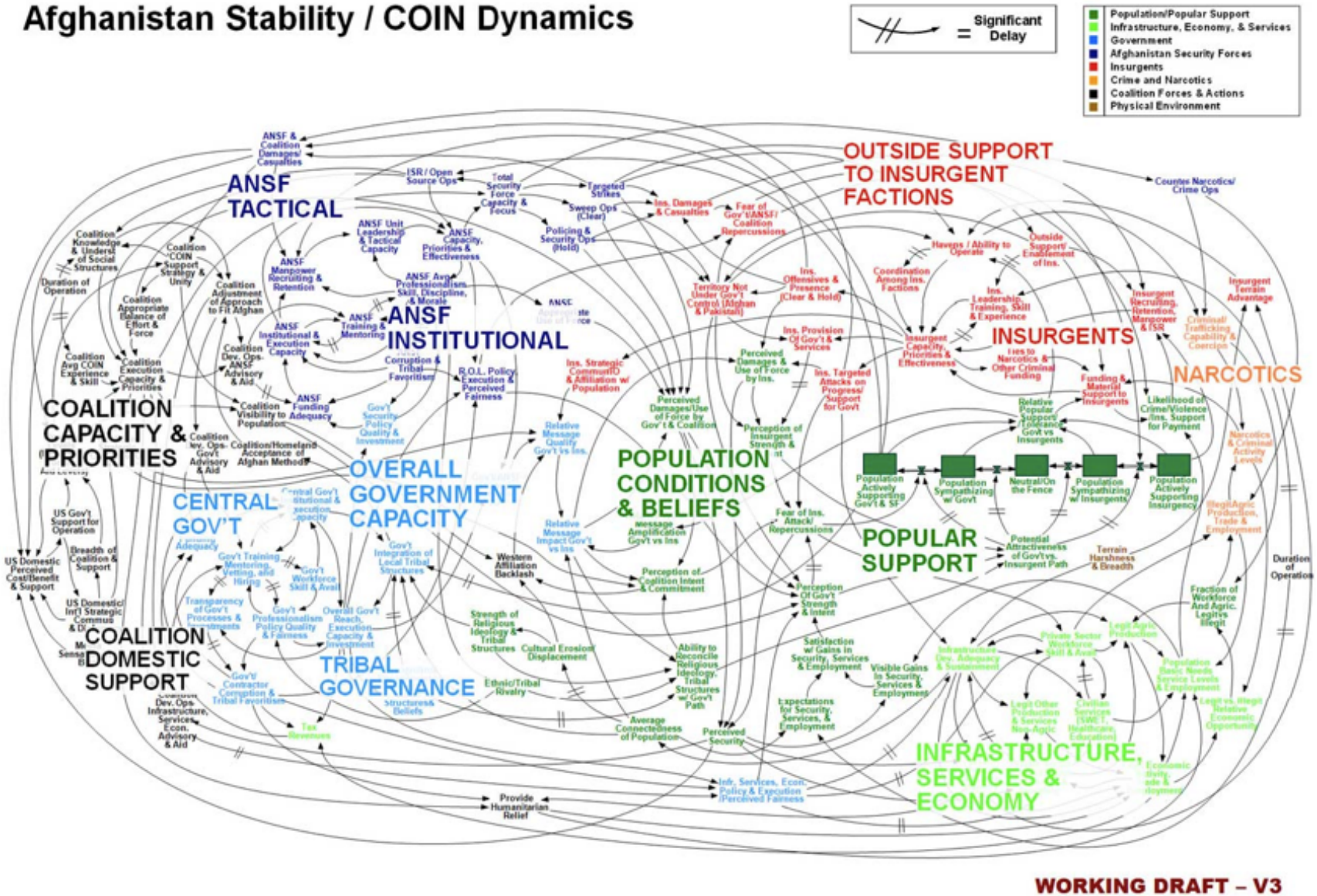
- Specification error: no model of a complex, open system can contain all of the relevant variables;
- Measurement error: with very few exceptions, variables will contain some measurement error
  - presupposing there is even agreement on what the "correct" measurement is in an ideal setting;
  - Predictive accuracy is limited by the square root of measurement error: in a bivariate model if your reliability is 80%, your accuracy can't be more than 90%
  - This biases the coefficient estimates as well as the predictions
- Quasi-random structural error: Complex and chaotic deterministic systems behave as if they were random under at least some parameter combinations .
  Chaotic behavior can occur in equations as simple as
  $x_{t+1} = ax_t{}^2 + bx_t$

Irreducible sources of error-2

- Rational randomness such as that predicted by mixed strategies in zero-sum games
- Arational randomness attributable to free-will
  - Rule-of-thumb from our rat-running colleagues:
    "A genetically standardized experimental animal, subjected to carefully controlled stimuli in a laboratory setting, will do whatever it wants."
- Effective policy response:
  in at least some instances organizations will have taken steps to head off a crisis that would have otherwise occurred.
- The effects of natural phenomenon
  - the 2004 Indian Ocean tsunami dramatically reduced violence in the long-running conflict in Aceh

(Tetlock (2013) independently has an almost identical list of the irreducible sources of error.)

Open, complex systems



Balancing factors which make behavior predictable

- Individual preferences and expectations, which tend to change very slowly
- Organizational and bureaucratic rules and norms
- Constraints of mass mobilization strategies
- Structural constraints:
  the Maldives will not respond to climate-induced sea level rise by building a naval fleet to conquer Singapore.
- Choices and strategies at Nash equilibrium points
- Autoregression (more a result than a cause)
- Network and contagion effects (same)

"History doesn't repeat itself but it rhymes"
Mark Twain (also occasionally attributed to Friedrich Nietzsche)

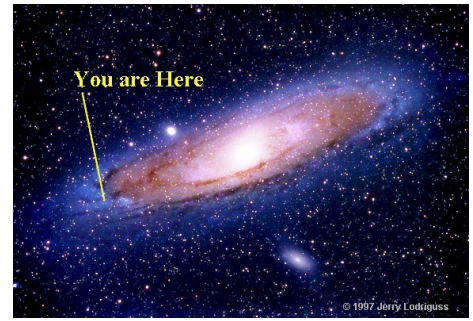
Paradox of political prediction

Political behaviors are generally highly incremental and vary little from day to day, or even century to century (Putnam).

Nonetheless, we *perceive* politics as very unpredictable because we focus on the unexpected (Kahneman).

Consequently the only “interesting” forecasts are those which are least characteristic of the system as a whole. However, only some of those changes are actually predictable.

Finding a non-trivial forecast



- Too frequent: prediction is obvious without technical assistance
- Too infrequent: prediction may be correct, but the event is so infrequent that
  - The prediction is irrelevant to policy
  - Calibration can be very tricky
  - Accuracy of the model is difficult to assess
- "Just right": these are situations where typical human accuracy is likely to be flawed, and consequently these could have a high payoff, but there are not very many of them.

Differing attitudes towards error

Geography:

- Progress consists of ever more accurate data

Political science:

- Trust nothing—everything has error, just control for the systematic biases

From presentation to a geospatial intelligence conference, where I suspect it had zero impact. . .

The Political Science View of the World

2 October 2012: Israeli Prime Minister Netanyahu reported to be pursuing increased sanctions against Iran

- He has changed his mind and thinks sanctions are an effective approach
  (face value)
- He has concluded Obama will be re-elected
  (international considerations)
- The Israeli military finally persuaded him an attack was a bad idea
  (domestic considerations)
- Israel is going to attack Iran in the near future
  (deception strategy)

The Political Science View of the World

2 October 2012: Israeli Prime Minister Netanyahu reported to be pursuing increased sanctions against Iran

- He has changed his mind and thinks sanctions are an effective approach
  (face value)
- He has concluded Obama will be re-elected
  (international considerations)
- The Israeli military finally persuaded him an attack was a bad idea
  (domestic considerations)
- Israel is going to attack Iran in the near future
  (deception strategy)

The Political Science View of the World

2 October 2012: Israeli Prime Minister Netanyahu reported to be pursuing increased sanctions against Iran

- He has changed his mind and thinks sanctions are an effective approach
  (face value)
- He has concluded Obama will be re-elected
  (international considerations)
- The Israeli military finally persuaded him an attack was a bad idea
  (domestic considerations)
- Israel is going to attack Iran in the near future
  (deception strategy)

(presumably apocryphal) exchange at the Congress of Vienna

- Aide: Your excellency, the Russian Ambassador has just died!
- Prince von Mitternich: Fascinating...now, what are his intentions?

Differing attitudes towards error

Geography:

- Progress consists of ever more accurate data

Political science:

- Trust nothing—everything has error, just control for the systematic biases

Machine learning::

- it is what it is: goal is improving prediction

Statistics::

- signal to noise: Perfect is the enemy of "good enough"
- mathematically approximate the characteristics of the error
- Taleb, Mandelbrot: don’t be a Gaussian in a power-law world

Models matter

Arab Spring is an unprecedented product of the new social media

- Model used by Chinese censors of NSM: King, Peng, Roberts 2012
- Next likely candidates: Africa

Arab Spring is an example of an instability contagion/diffusion process

- Eastern Europe 1989-1991, OECD 1968, CSA 1859-1861, Europe 1848, Latin America 1820-1828
- Next likely candidates: Central Asia

Arab Spring is a black swan

- There is no point in modeling black swans, you instead build systems robust against them

Statistical and modeling challenges

Rare events

- Incorporate much longer historical time lines?—Schelling used Caesar's *Gallic Wars* to analyze nuclear deterrence
- New approaches made possible by computational advances

Analysis of event sequences, which are not a standard data type

- There are, however, a large number of available methods, and it is just possible that these will work with very large data sets such as GDELT
- This possibility will be discussed in detail in Lecture 5

Causality

- Oxford *Handbook of Causation* is 800 pages long

Integration of qualitative and qualitative/subject-matter-expert (SME) information

- Bayesian approaches using prior probabilities are promising but to date they have not really been used

Pournelle's Law:
No task is so virtuous that it will not attract idiots

- Need to establish with the media and policy-makers that not every forecast, ~~even~~ especially those made using "Big Data" methods, is scientifically valid
  - It took the survey research community about thirty to forty years to establish professional credibility, though they have largely succeeded
- Conveying limitations of the methods against the hyper-confidence of pundits and individuals with secret models
  - Limitations of the data sources
  - Limitations of the data coding, particularly automated coding
  - Limitations of the model estimation
  - Limitations of probabilistic forecasts, particularly for rare events, even when the models are correct

Critical case: studies of climate change and conflict. As Pinker and Goldstein noted, people want to hear simple scary answers.

Ethical concerns

- Thus far, we've generally had the luxury of no one paying attention to any of our predictions : what if governments do start paying attention?
  - "Policy relevant forecast interval" is around 6 to 24 months
  - USAID/FAO famine forecasting model
  - It is *possible* that our models could become less accurate because crises are being averted, but I don't see that happening any time soon.
- Difficulties in getting *anyone*, including experts (see Kahneman, Tetlock), to correctly interpret probabilistic forecasts
- Possible impact on sources
  - Local collaborators
  - Journalists (cf. Mexico)
  - NGOs to the extent we are using their information

Thank you

Email: schrodt735@gmail.com

Slides: http://eventdata.parusanalytics.com/presentations.html

Forecasting papers:
http://eventdata.parusanalytics.com/papers.html