Predictive Policing and the ETAS Model

National Association of Criminal Defense Lawyers
San Francisco, CA

Philip B. Stark www.stat.berkeley.edu/~stark

11 February 2019

Department of Statistics, University of California, Berkeley

- How does predictive policing work?
- Does ETAS work in seismology?
- Is PredPol software special / worth the cost?
- Is PredPol ‘fair’?

Statistical models for crime: metaphor, not criminology

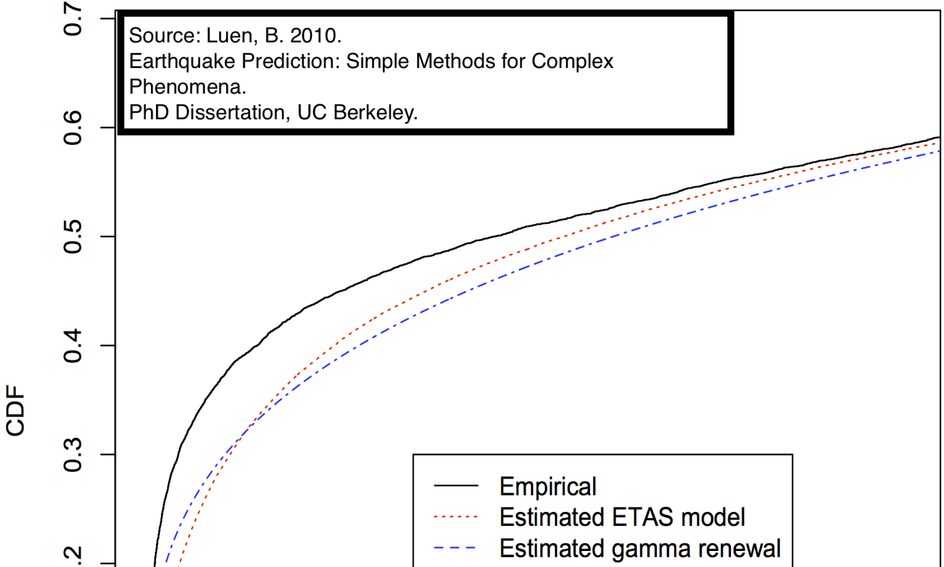
- Phenomenology: some crimes occur in clusters
- Crime occurs "as if" in a casino game whose rules are embodied in some mathematical model known to the person who wrote the software.
- Like saying that there is a deck of crime cards for each "block"
  - Deck contains some blank cards and cards with various crimes on them
  - In a given block, in every time interval, a card is dealt from the deck
  - If card is blank, no crime
  - Otherwise, there is a crime of the kind on the card
  - In ETAS model, if you draw a card with a crime, deck gets extra cards with that crime (extras gradually removed in successive draws)

Different models make different assumptions about # cards of each type there are, the shuffling, whether drawn cards are returned to the deck, etc.

ETAS: Draw w/ replacement. Crime begets additional crime of the same kind, in the same block. Each crime has zero or one "parent."

Example of *linear marked Hawkes process*

Inter-event times of M>=3 earthquakes, Southern California



ETAS Parameter estimates & simulations

- Often unphysical for real data (each event expected to have infinitely many children)
- Simulations can have burn-in times of order 10^5y

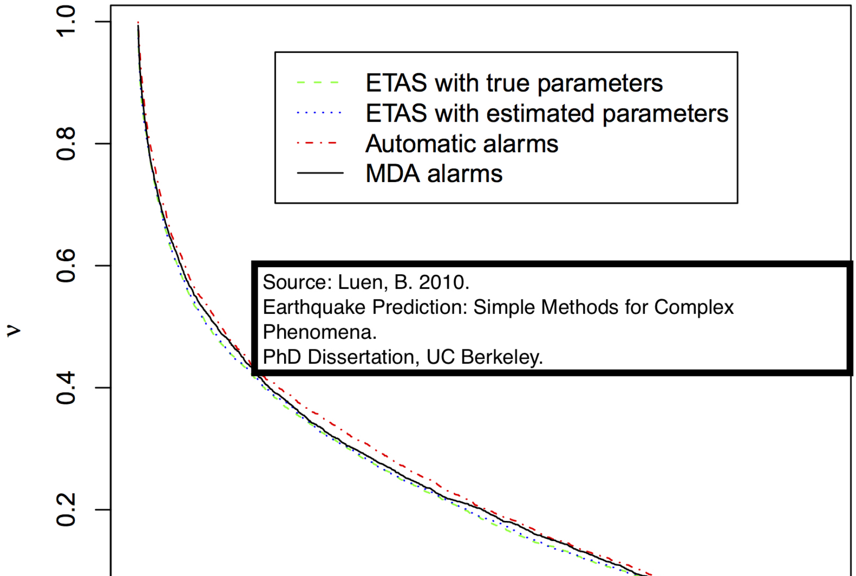
Sornette & Utkin, 2009. *Phys. Rev. E 79*, 061110

- Simulate ETAS seismicity
- Use ETAS to classify event as background or child (aftershock)
- Unreliable
  - estimated rates of exogenous events suffer from large errors
  - branching ratio badly estimated in general
  - high level of randomness together with the long memory makes the stochastic reconstruction of trees of ancestry and the estimation of the key parameters perhaps intrinsically unreliable

Prediction: Automatic Alarms, MDA, & ETAS

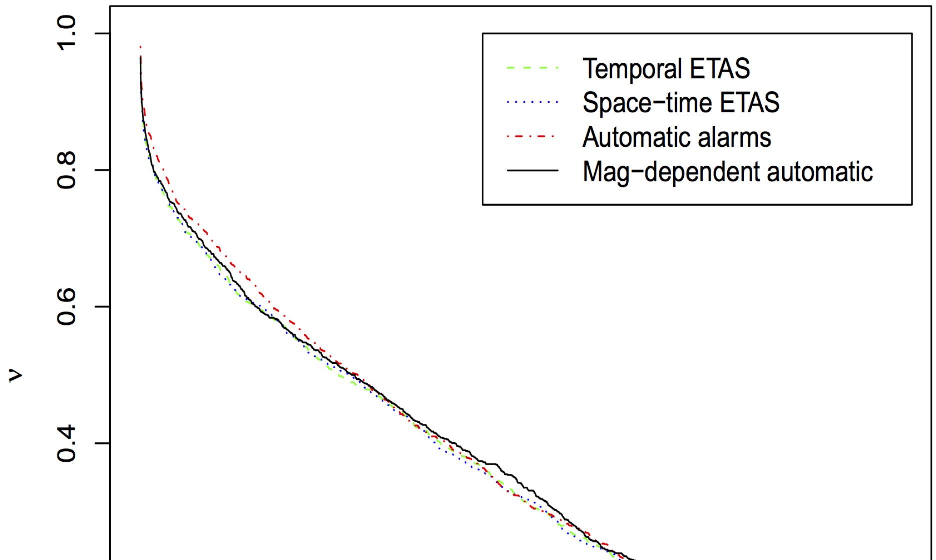
- Automatic alarm: after every event with $M > \mu$, start an alarm of duration τ
- Magnitude-dependent automatic alarm (MDA): after every event with $M > \mu$, start an alarm of duration τu^M

Error diagrams for predictors of ETAS simulation



Source: Luen, B. 2010.
Earthquake Prediction: Simple Methods for Complex Phenomena.
PhD Dissertation, UC Berkeley.

Error diagrams for predictions of Southern Californian seismicity



Randomized Controlled Field Trials of Predictive Policing

G. O. Mohler, M. B. Short, Sean Malinowski, Mark Johnson, G. E. Tita, Andrea L. Bertozzi, and P. J. Brantingham

The concentration of police resources in stable crime hotspots has proven effective in reducing crime, but the extent to which police can disrupt dynamically changing crime hotspots is unknown. Police must be able to anticipate the future location of dynamic hotspots to disrupt them. Here we report results of two randomized controlled trials of near real-time epidemic-type aftershock sequence (ETAS) crime forecasting, one trial within three divisions of the Los Angeles Police Department and the other trial within two divisions of the Kent Police Department (United Kingdom). We investigate the extent to which (i) ETAS models of short-term crime risk outperform existing best practice of hotspot maps produced by dedicated crime analysts, (ii) police officers in the field can dynamically patrol predicted hotspots given limited resources, and (iii) crime can be reduced by predictive policing algorithms under realistic law enforcement resource constraints. While previous hotspot policing experiments fix treatment and control hotspots throughout the experimental period, we use a novel experimental design to allow treatment and control hotspots to change dynamically over the course of the experiment. Our results show that ETAS models predict 1.4–2.2 times as much crime compared to a dedicated crime analyst using existing criminal intelligence and hotspot mapping practice. Police patrols using ETAS forecasts led to an average 7.4% reduction in crime volume as a function of patrol time, whereas patrols based upon analyst predictions showed no significant effect. Dynamic police patrol in response to ETAS crime forecasts can disrupt opportunities for crime and lead to real crime reductions.

KEY WORDS: Crime; Experimental methods; Machine learning; Point processes; Policing dosage.

Figure 4: Mohler et al.

2.2 Epidemic-Type Aftershock Sequence Model for Crime Prediction

Building on a foundation of reaction-diffusion models of crime (Short et al. 2010), we treat the dynamic occurrence of crime as a continuous time, discrete space ETAS point process (Marsan and Lengline 2008; Mohler et al. 2011; Mohler 2014). Policing areas were first discretized into 150 × 150 m square boxes. The conditional intensity, or probabilistic rate $\lambda_n(t)$ of events in box n at time t was determined by

$$\lambda_n(t) = \mu_n + \sum_{t_n^i < t} \theta \omega e^{-\omega(t - t_n^i)}, \qquad (1)$$

Figure 5: Mohler et al. re ETAS

The EM algorithm can be intuitively understood by viewing the ETAS model as a branching process (Mohler et al. 2011). First-generation events occur according to a Poisson process with constant rate μ. Events (from all generations) each give birth to N direct offspring events, where N is a Poisson random variable with parameter θ. As events occur, the rate of crime increases locally in space, leading to a contagious sequence of "aftershock" crimes (Mohler et al. 2011) that eventually dies

Figure 6: Mohler et al. re ETAS

The emphasis in Los Angeles was reversed. The Los Angeles Police Department follows a COMPSTAT (Walsh 2001) policing model focused on the analysis of 7-day crime maps supplemented with ad hoc street-level intelligence. The LAPD analysts emphasized small clusters of recent crimes as signaling emerging problems in need of short-term response to interrupt the

Figure 7: Mohler et al. re analyst

STAT 157 Predictive Policing Paper

Ellen Kulinsky
SID 26126845

Jong Ha Lee
SID 25344865

Evan Limanto
SID 25557086

Vaibhav Ramamoorthy
SID 25539060

Arun Ramamurthy
SID 26440691

December 15, 2017

Abstract

Since its inception in 2011 from the LAPD, predictive policing – the usage of mathematical and analytical techniques in law enforcement to identify potential criminal activity – has pro-

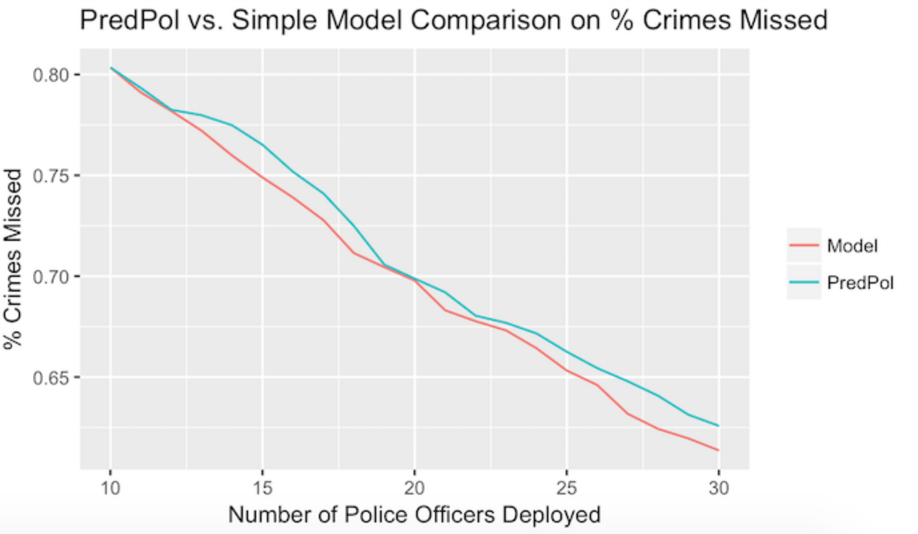


Figure 9: Kulinsky et al. 2017 ROC

IN DETAIL

To predict and serve?

Predictive policing systems are used increasingly by law enforcement to try to prevent crime before it occurs. But what happens when these systems are trained using biased data? **Kristian Lum** and **William Isaac** consider the evidence – and the social consequences

Discussion

We have demonstrated that predictive policing of drug crimes results in increasingly disproportionate policing of historically over-policed communities. Over-policing imposes real costs on these communities. Increased police scrutiny and surveillance have been linked to worsening mental and physical health;[10,11] and, in the extreme, additional police contact will create additional opportunities for police violence in over-policed areas.[12] When the costs of policing are disproportionate to the level of crime, this amounts to discriminatory policy.

Figure 11: Lum Isaac 2016

We find that rather than correcting for the apparent biases in the police data, the model reinforces these biases. The locations that are flagged for targeted policing are those that were, by our estimates, already over-represented in the historical police data. Figure 2(b) shows the percentage of the population experiencing targeted policing for drug crimes broken down by race. Using PredPol in Oakland, black people would be targeted by predictive policing at roughly twice the rate of whites. Individuals classified as a race other than white or black would receive targeted policing at a rate 1.5 times that of whites. This is in contrast to the estimated pattern of drug use by race, shown in Figure 2(c), where drug use is roughly equivalent across racial classifications. We find similar results when analysing the rate of targeted policing by income group, with low-income households experiencing targeted policing at disproportionately high rates.

Conclusions

- Some crimes cluster in time and space
- ETAS tries to exploit clustering
  - based on heuristics & metaphors, not criminology
  - "borrows strength" from seismology–where it doesn't work very well
  - simpler/cheaper methods may do just as well
  - comparisons in Mohler et al. 2018
- PredPol exacerbates policing biases in "training" data
  - reporting and enforcement uneven
  - screen of "objectivity"
  - Mohler et al. study not statistically convincing
  - much hype

References

- Geller, R.J., F. Mulargia, and P.B. Stark, 2015. Why We Need a New Paradigm of Earthquake Occurrence, in *Subduction Dynamics: From Mantle Flow to Mega Disasters*, AGU, https://doi.org/10.1002/9781118888865.ch10
- Kulinsky, E., J.H. Lee, E.W. Limanto, V. Ramamoorthy, and A. Ramamurthy, 2017. Stat 157 Predictive Policing Paper, https://github.com/arun-ramamurthy/pred-pol/blob/s157/presentation/STAT_157_Predictive_Policing_Paper.pdf
- Luen, B., 2010. *Earthquake prediction: Simple methods for complex phenomena*, Ph.D. dissertation, University of California, Berkeley http://digitalassets.lib.berkeley.edu/etd/ucb/text/Luen_berkeley_0028E_10884.pdf

- Luen, B., and P.B. Stark, 2008. Testing earthquake predictions, in *Probability and Statistics: Essays in Honor of David A. Freedman*, 302–315, Institute of Mathematical Statistics, DOI: 10.1214/193940307000000509
- Lum, K., and W. Isaac, 2016. To Predict and Serve, *Significance Magazine*, https://doi.org/10.1111/j.1740-9713.2016.00960.x
- Mohler, G.O., M.B. Short, S. Malinowski, M. Johnson, G.E. Tita, A.L. Bertozzi, and P.J. Brantingham, 2018. Randomized controlled field trials of predictive policing, *JASA, 110*, DOI: 10.1080/01621459.2015.1077710
- Mulargia, F., Stark, P.B., and R.J. Geller, 2017. Why is Probabilistic Seismic Hazard Analysis still used? *Physics of the Earth and Planetary Interiors, 264*, 63–75 https://www.sciencedirect.com/science/article/pii/S0031920116303016