Statistical framework for decision making in mine action

Jan Larsen
Intelligent Signal Processing Group
Informatics and Mathematical Modelling
Technical University of Denmark
jl@imm.dtu.dk, www.imm.dtu.dk/~jl

How do we construct a reliable detector?

- Empirical method: systematic acquisition of knowledge which is used to build a mathematical model to generate reliable results in real use cases
- Specifying the relevant scenarios and performance measures – end user involvement is crucial!!!
- Cross-disciplinary R&D involving very competences

Knowledge acquisition

Physical modeling

- Study physical properties and mechanism of the environment and sensors
- Describe the knowledge as a mathematical model

Statistical modeling

Require real world related data

Use data to learn e.g. the relation between the sensor reading and the presence/absence of explosives

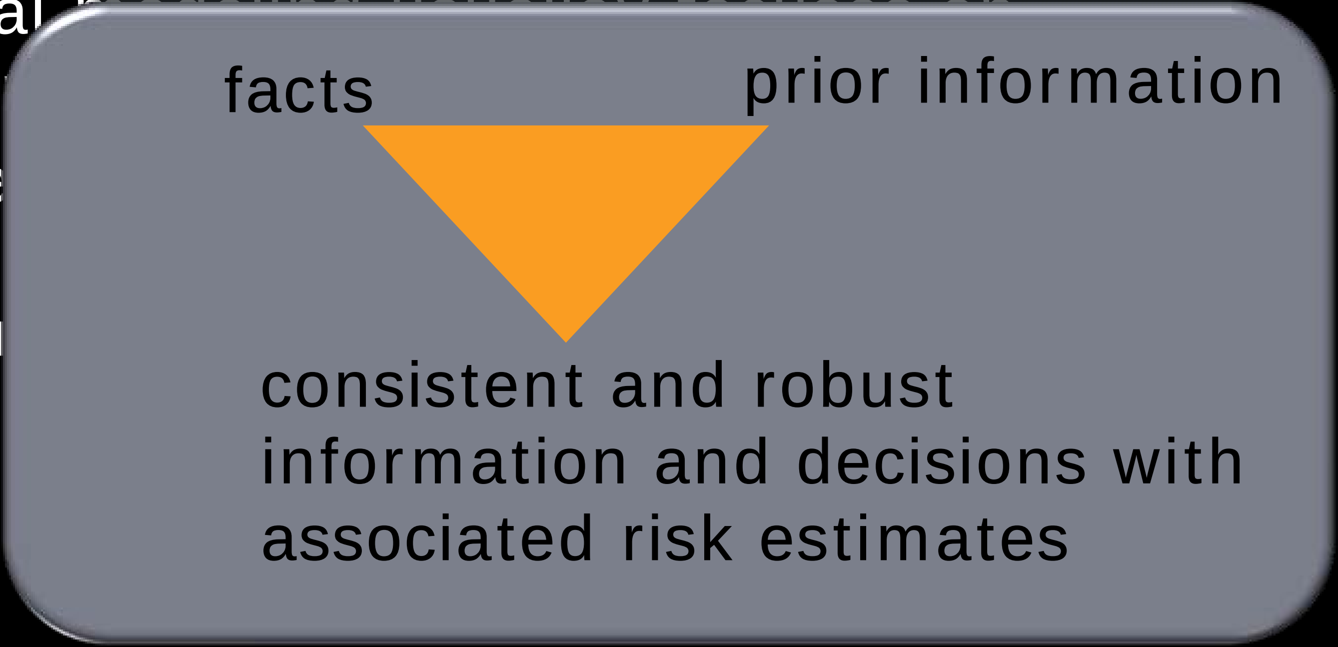
Why do we need statistical models and machine learning?

- Mine action is influenced by many uncertain factors
- The goals of mine action depends on difficult socio-economic and political considerations and constraints are to be built in

Scientist are born sceptical: they don’t believe facts unless they see them often enough

Why do we need statistical models and machine learning?

- statistical modeling is the principled framework to handle uncertainty and complexity
- statistical
  identifyi
- machine
  collectio
  new situ



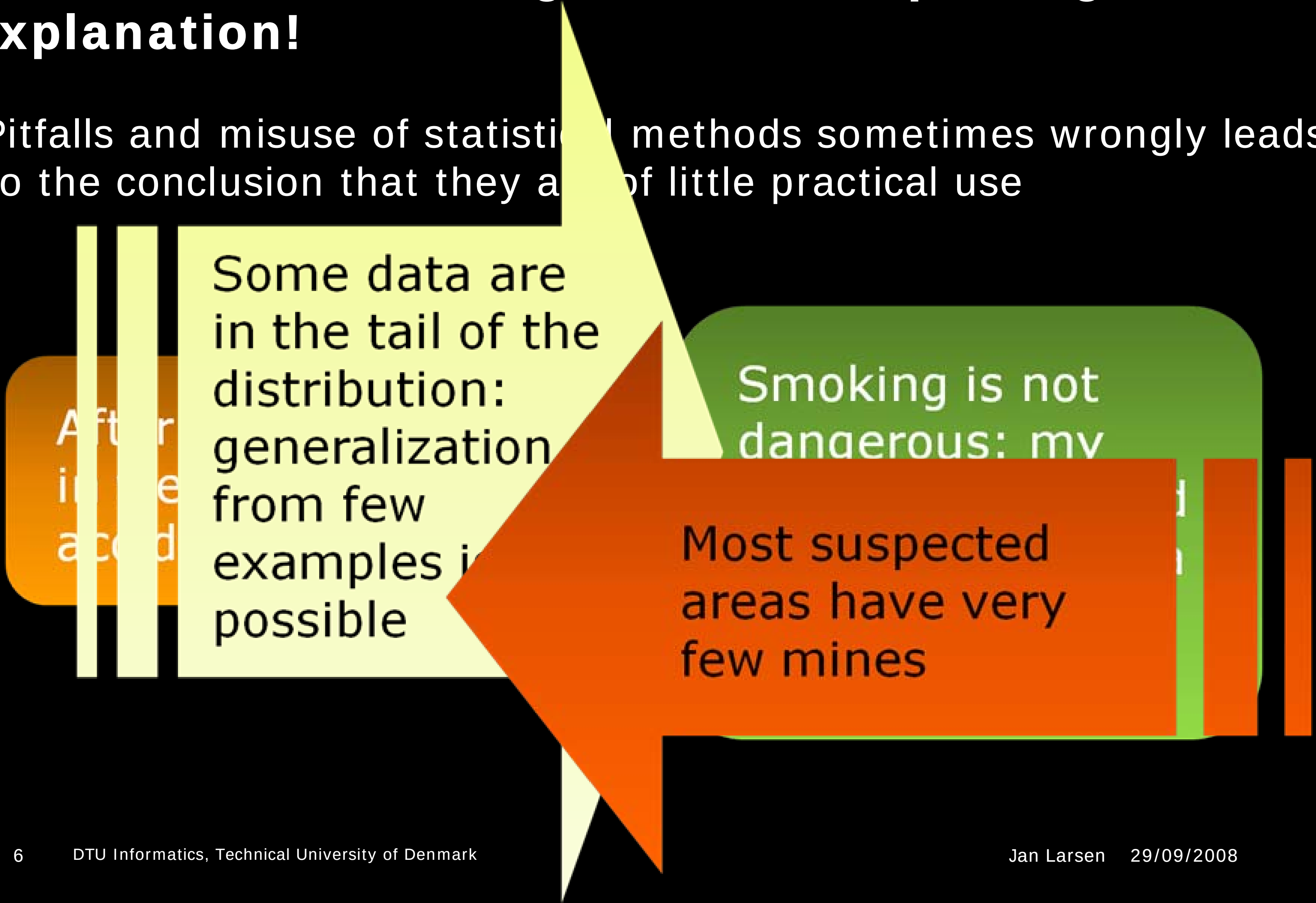
facts

prior information

consistent and robust information and decisions with associated risk estimates

There is no such thing as facts to spoil a good explanation!

- Pitfalls and misuse of statisti l methods sometimes wrongly leads to the conclusion that they a of little practical use



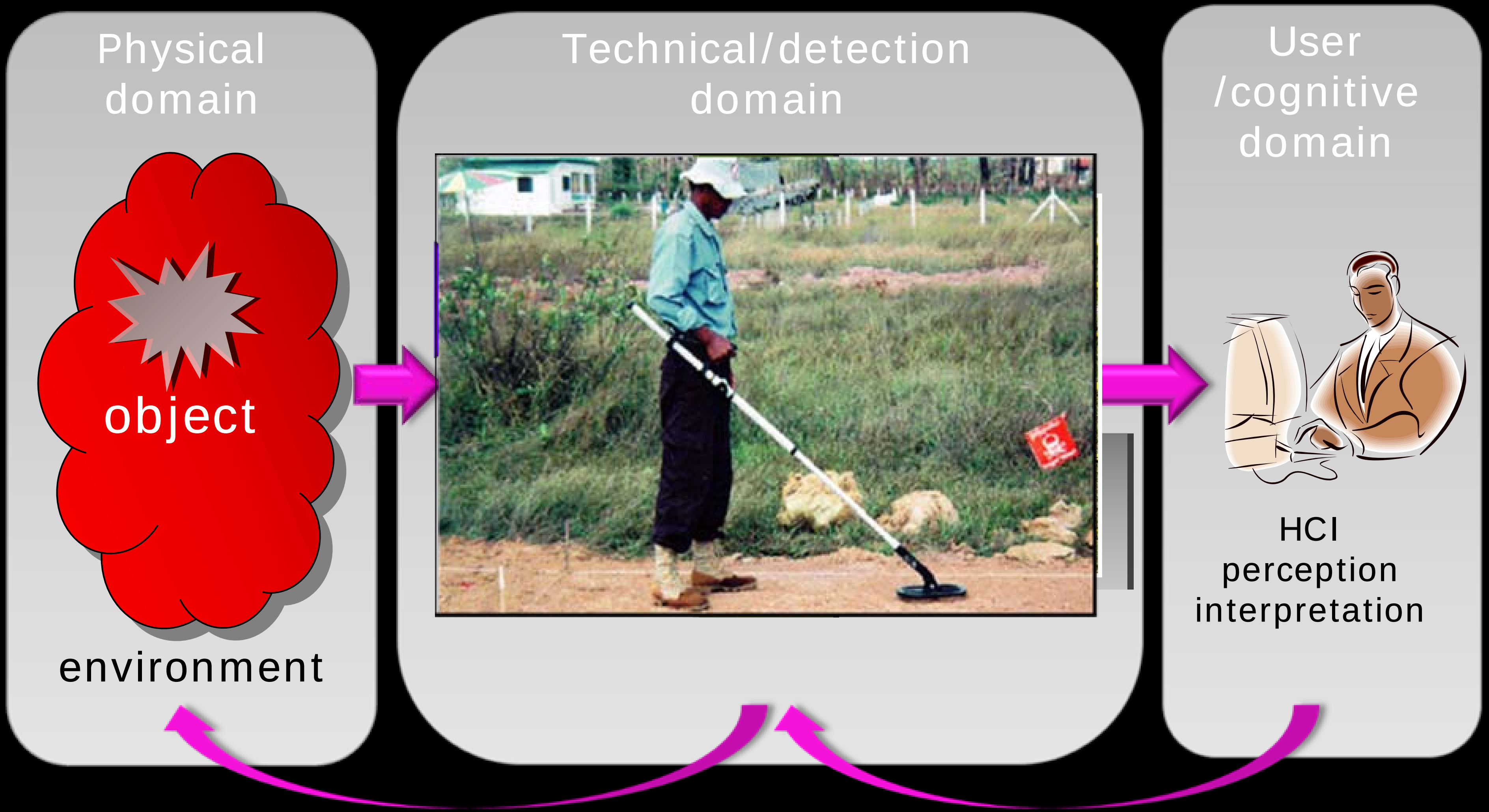
Aft r
i e
acc d

Some data are in the tail of the distribution: generalization from few examples i possible

Smoking is not dangerous: my

Most suspected areas have very few mines

Information processing pipeline



Physical domain

object

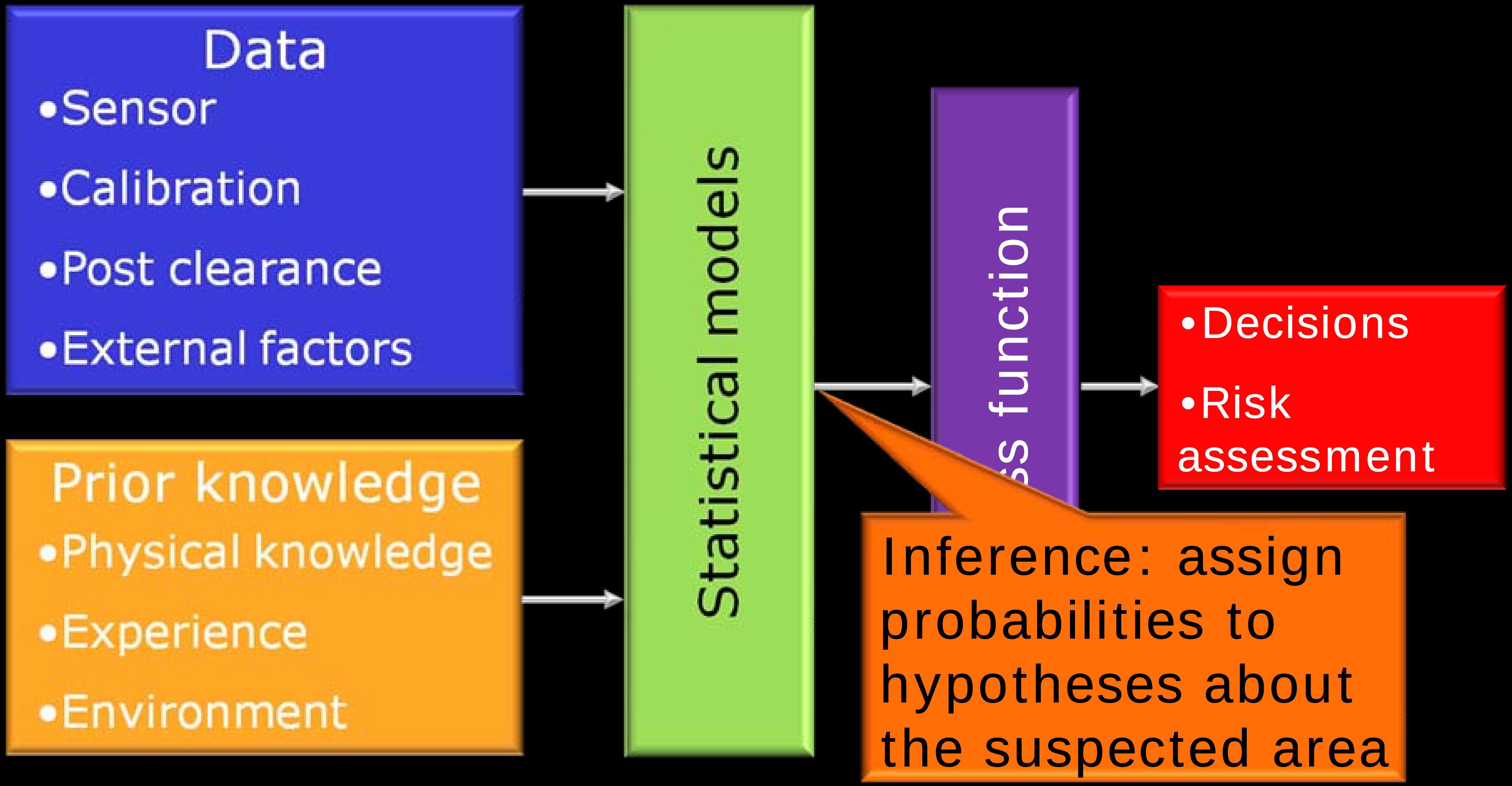
environment

Technical/detection domain

User /cognitive domain

HCI
perception
interpretation

The elements of statistical decision theory



Data
- Sensor
- Calibration
- Post clearance
- External factors

Prior knowledge
- Physical knowledge
- Experience
- Environment

Statistical models

s function

- Decisions
- Risk assessment

Inference: assign probabilities to hypotheses about the suspected area

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsense projects

Detecting a mine – tossing a coin

$$Frequency = \frac{\text{no of heads}}{\text{no of tosses}}$$

$$probability = frequency \text{ when infinitely many tosses}$$

On 99,6 % detection probability

$$\mathit{Frequency} = \frac{9960}{10000} = 99{,}60\%$$

One more (one less) count will change the frequency a lot!

Detection probability - tossing a coin

- N ndependent tosses number of
- y number of heads observed
- θ probability of heads

$$\hat{\theta} = \frac{y}{N}$$

$$P(y \mid \theta) = \mathrm{Binom}(\theta \mid N) = \binom{N}{y} \theta^{y} \theta^{N-y}$$

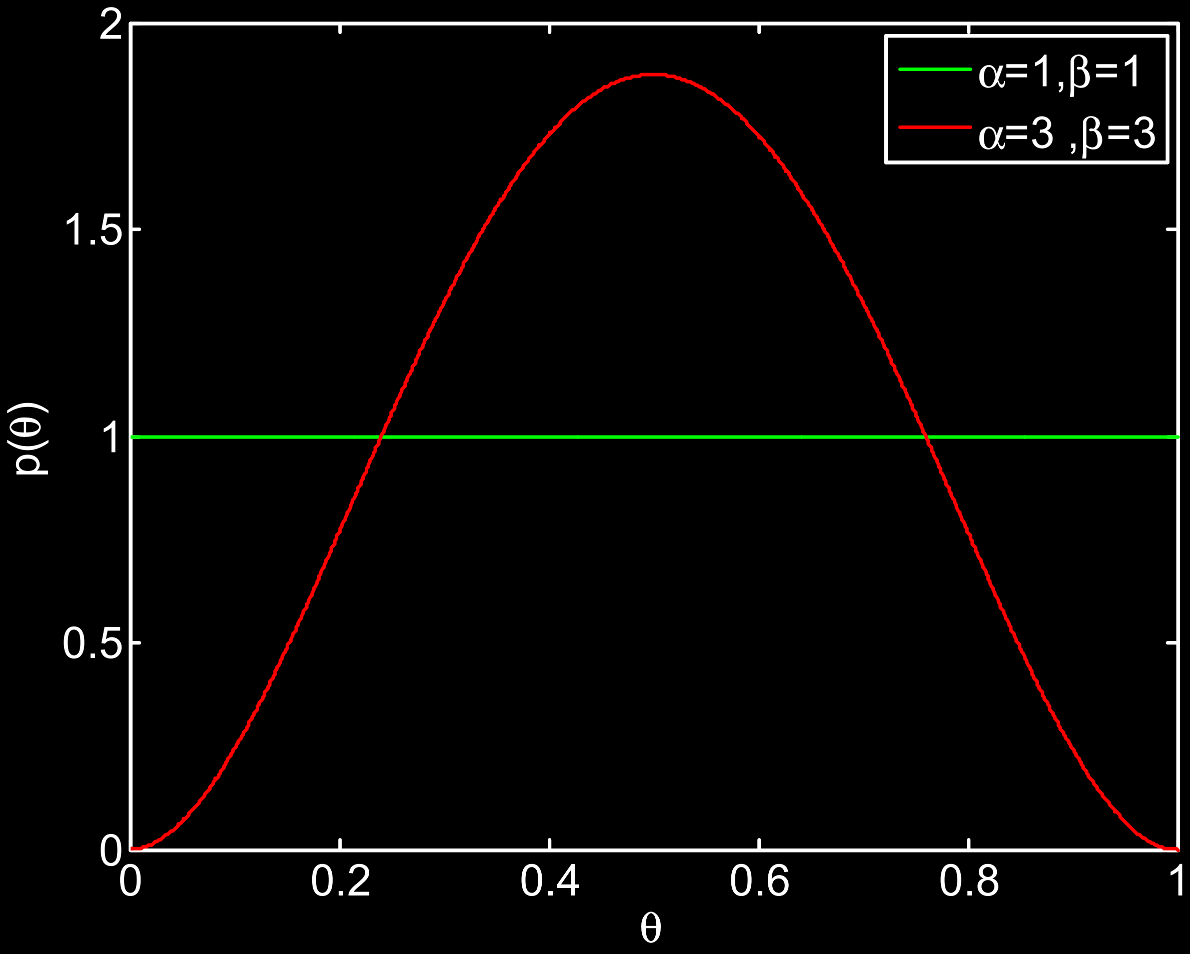
Data likelihood

Prior beliefs and opinions

- Prior 1: the fair coin: θ should be close to 0.5
- Prior 2: all values of θ are equally plausible

$$p(\theta) = Beta(\theta \mid \alpha, \beta)$$

Prior beliefs and opinions



Bayes rule: combining data likelihood and prior

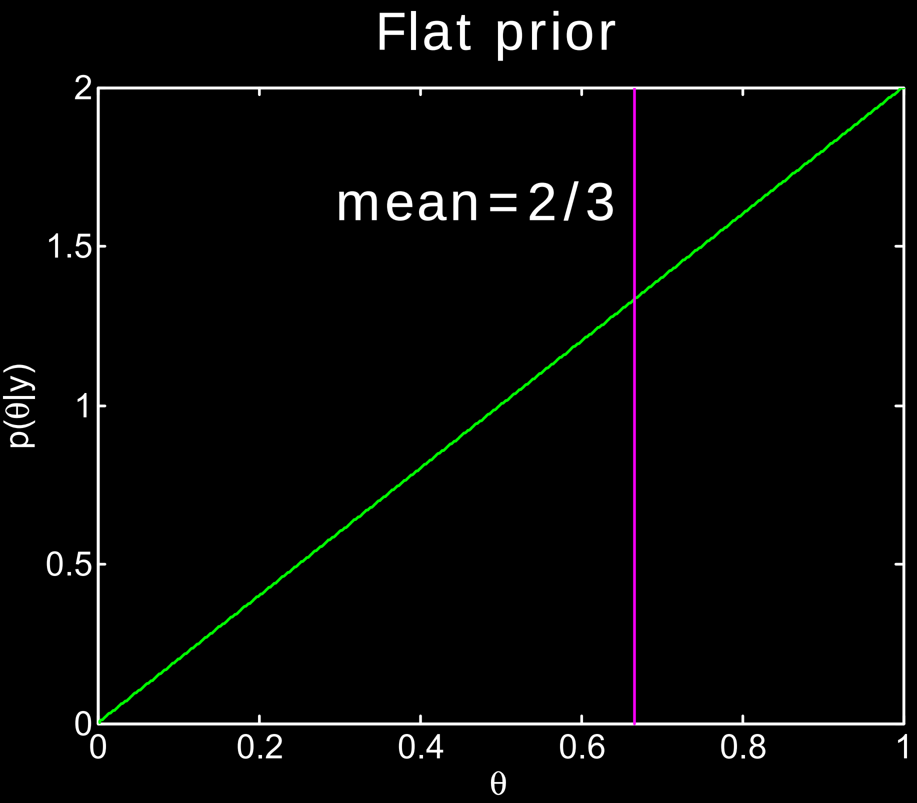
$$\underbrace{P(\theta \mid y)}_{\text{Posterior}} = \frac{\overbrace{P(y \mid \theta)}^{\text{Likelihood}}\,\overbrace{p(\theta)}^{\text{Prior}}}{P(y)}$$

$$P(\theta \mid y) = Beta(\theta \mid y + \alpha, \beta + n - y) \sim \theta^{y+\alpha}\theta^{n-y+\beta}$$

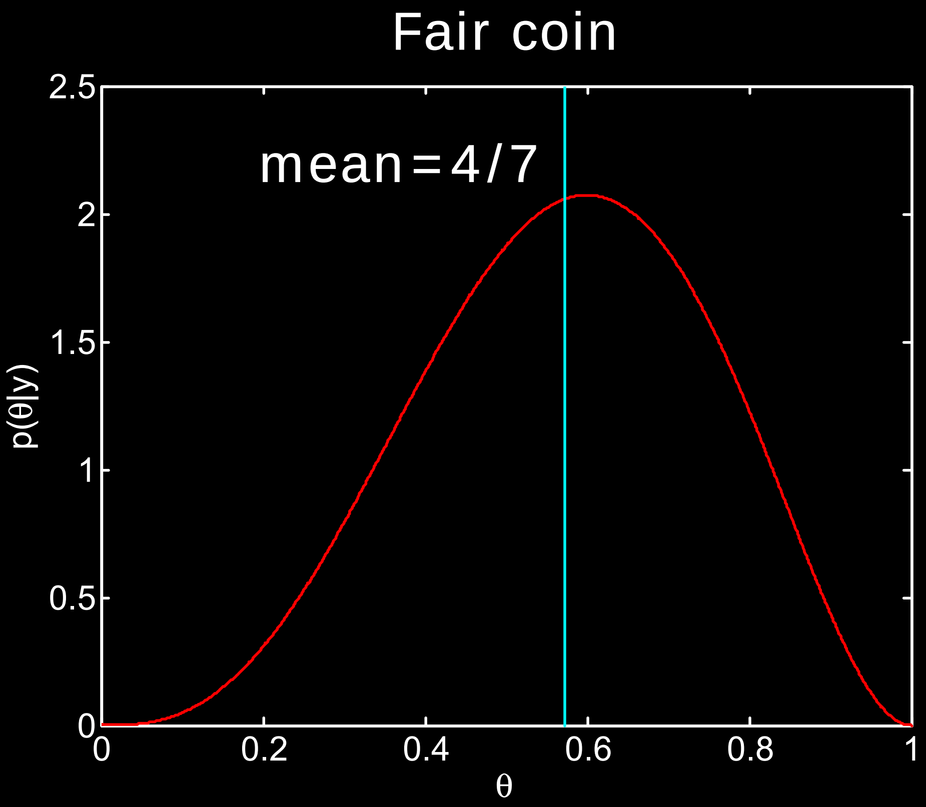
Posterior probability is also Beta

$$P(\theta \mid y) = Beta(\theta \mid y + \alpha, \beta + n - y) \sim \theta^{y+\alpha}\theta^{n-y+\beta}$$

Posteriors after observing one head



$Beta(\theta\mid 2,1)$



Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsens projects

What are the requirements for mine action risk

- Tolerable risk for individuals comparable to other natural risks
- As high cost efficiency as possible requires detailed risk management and control involving all partners (MAC, NGOs, commercial etc.)

99.6% detection is not an unrealistic requirement

but… today’s methods achieve at most 90% and are hard to evaluate!!!

GICHD and FFI are currently working on such methods [Håvard Bach, Ove Dullum NDRF SC2006]

A simple inference model – assigning probabilities to data

- The detection system provides the probability of detection a mine in a specific area: Prob(detect)
- The land area usage behavio of encounter: Prob(mine encou

Prob(casualty)=(1-Prob(d r)

For discussion of assumptions and involved factors see

“Risk Assessment of Minefields in HMA – a Bayesian Approach”

PhD Thesis, IMM/DTU 2005 by Jan Vistisen

A simple loss/risk model

- Minimize the number of casualties
- Under mild assumptions this equivalent to minimizing the probability of casualty

Maximum yearly footprint area in m^2

P(detection)	ρ : mine density (mines/km^2)				
	0.1	1	10	100	1000
0.996	25000	2500	250	25	2.5
0.9	1000	100	10	1	0.1

Prob(causality) = 10^{-5} per year

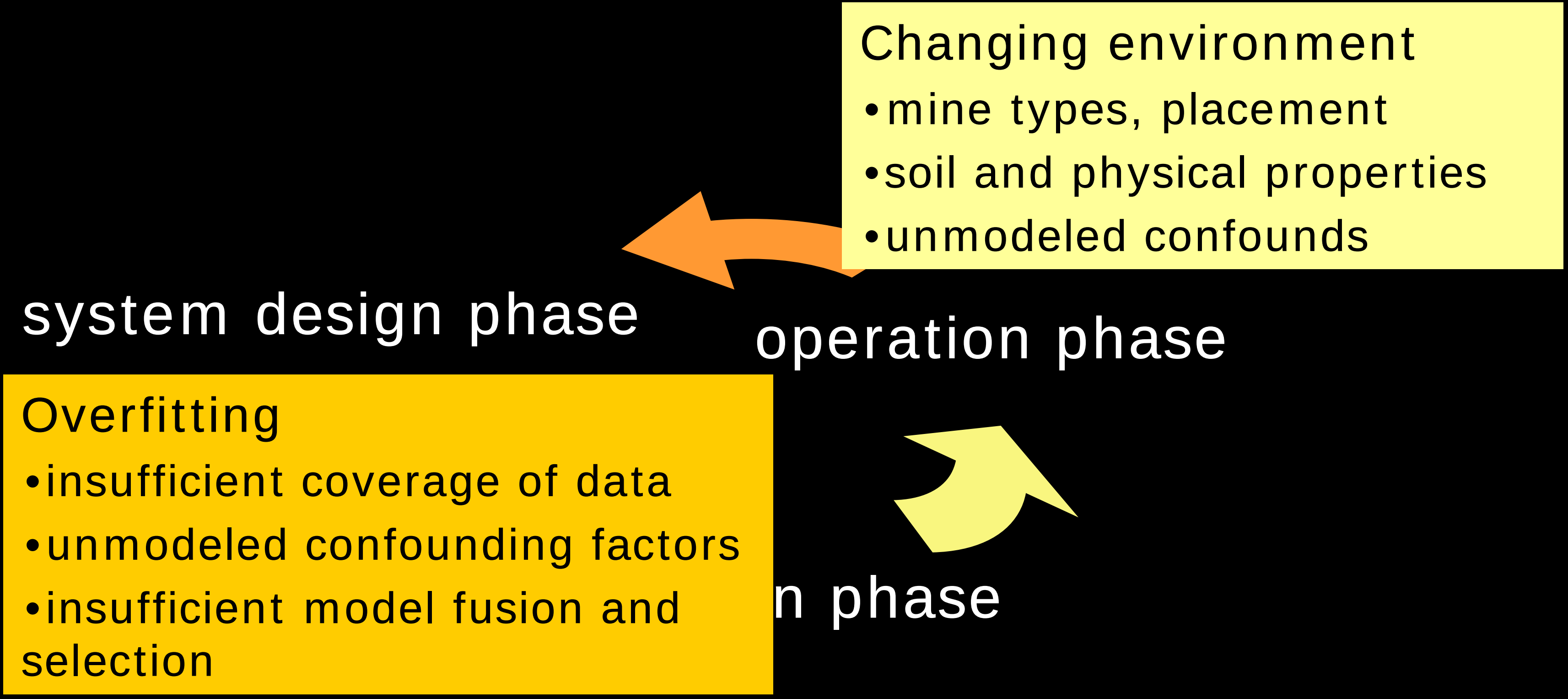
Reference: Bjarne Haugstad, FFI

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsense projects

Evaluation and testing in MA

- How do we assess the performance/detection probability?
- What is the confidence?



Changing environment

- mine types, placement
- soil and physical properties
- unmodeled confounds

system design phase

operation phase

n phase

Overfitting

- insufficient coverage of data
- unmodeled confounding factors
- insufficient model fusion and selection

Two types of error in detection of mines

Sensing error

The system does not sense properties of the mine object

⬇ decrease in detection probability

Decision error

The detector misinterprets the sensed signal

⬆ increase in false alarm rate

Example: metal detector

- Sensing error: the mine has low metal content
- Decision error: a piece of scrap metal was found

Example: mine detection dog

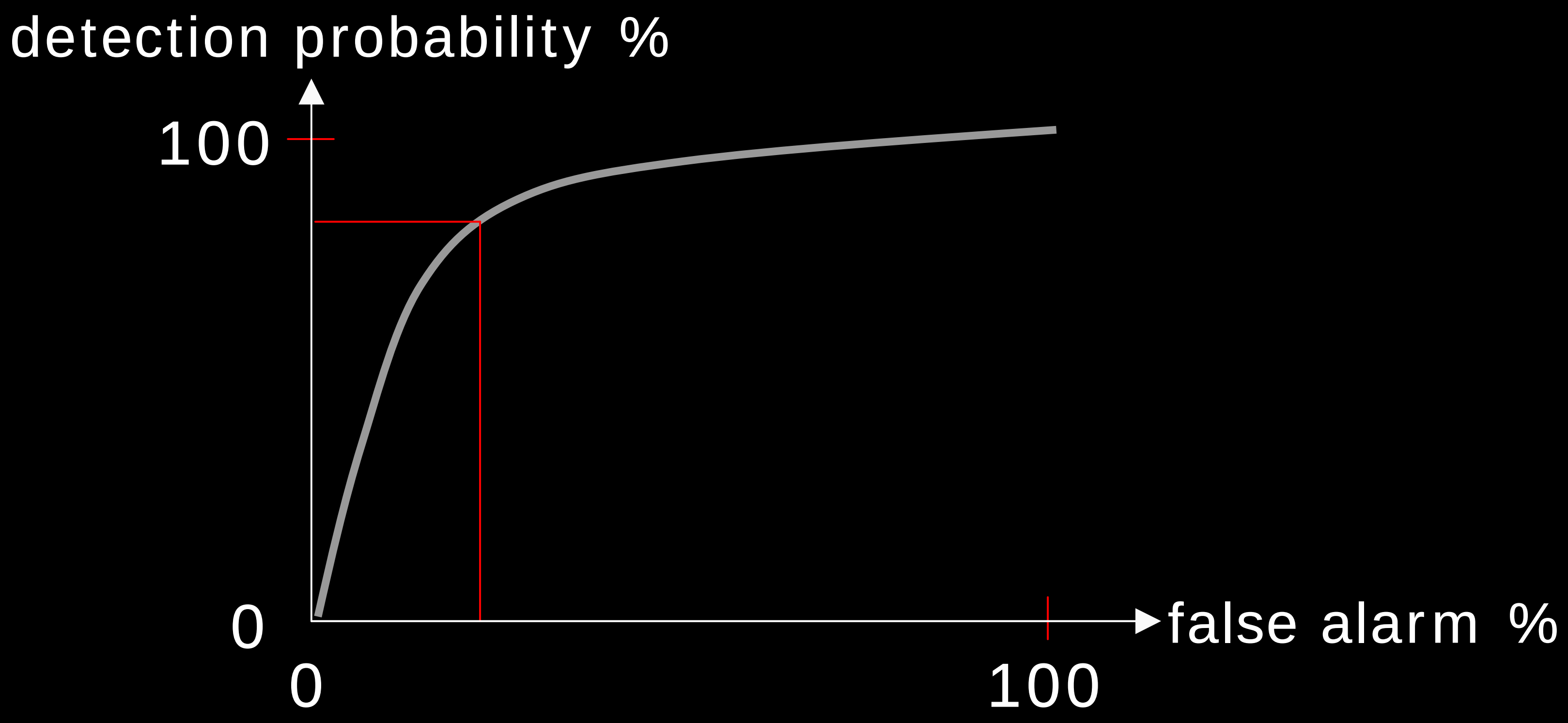
- Sensing error: the TNT leakage from the mine was too low
- Decision error: the dog handler misinterpreted the dogs indication

Confusion matrix in system design and test phase which should lead to certification

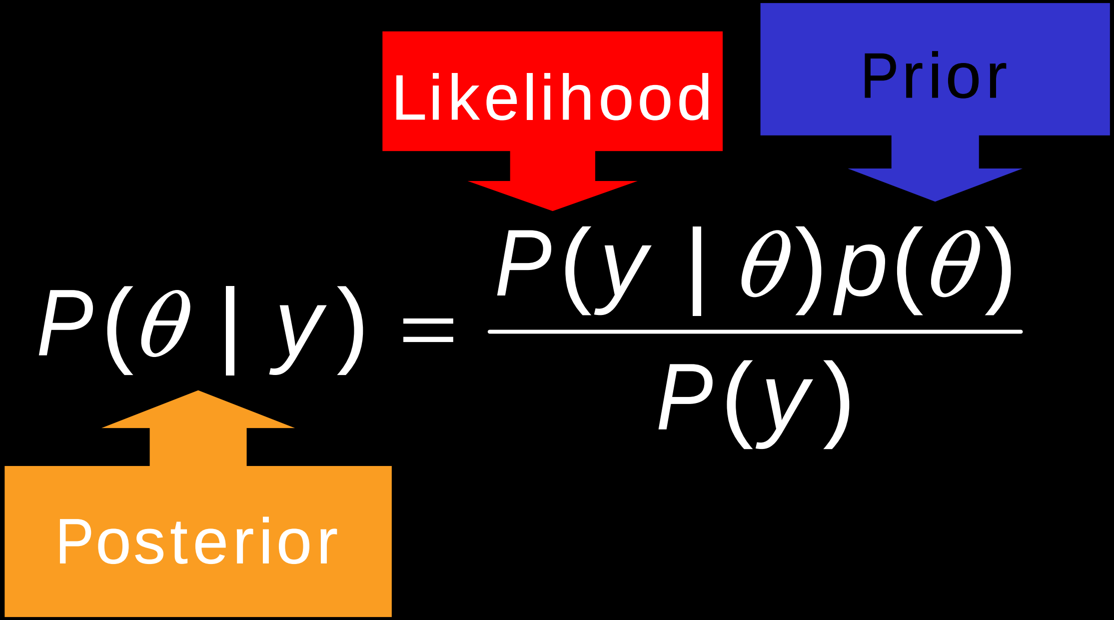
		True	
		yes	no
Estimated	yes	a	b
	no	c	d

- Detection probability (sensitivity): a/(a+c)
- False alarm: b/(a+b)
- False positive (specificity): b/(b+d)

Receiver operation characteristic (ROC)



Bayes rule: combining data likelihood and prior



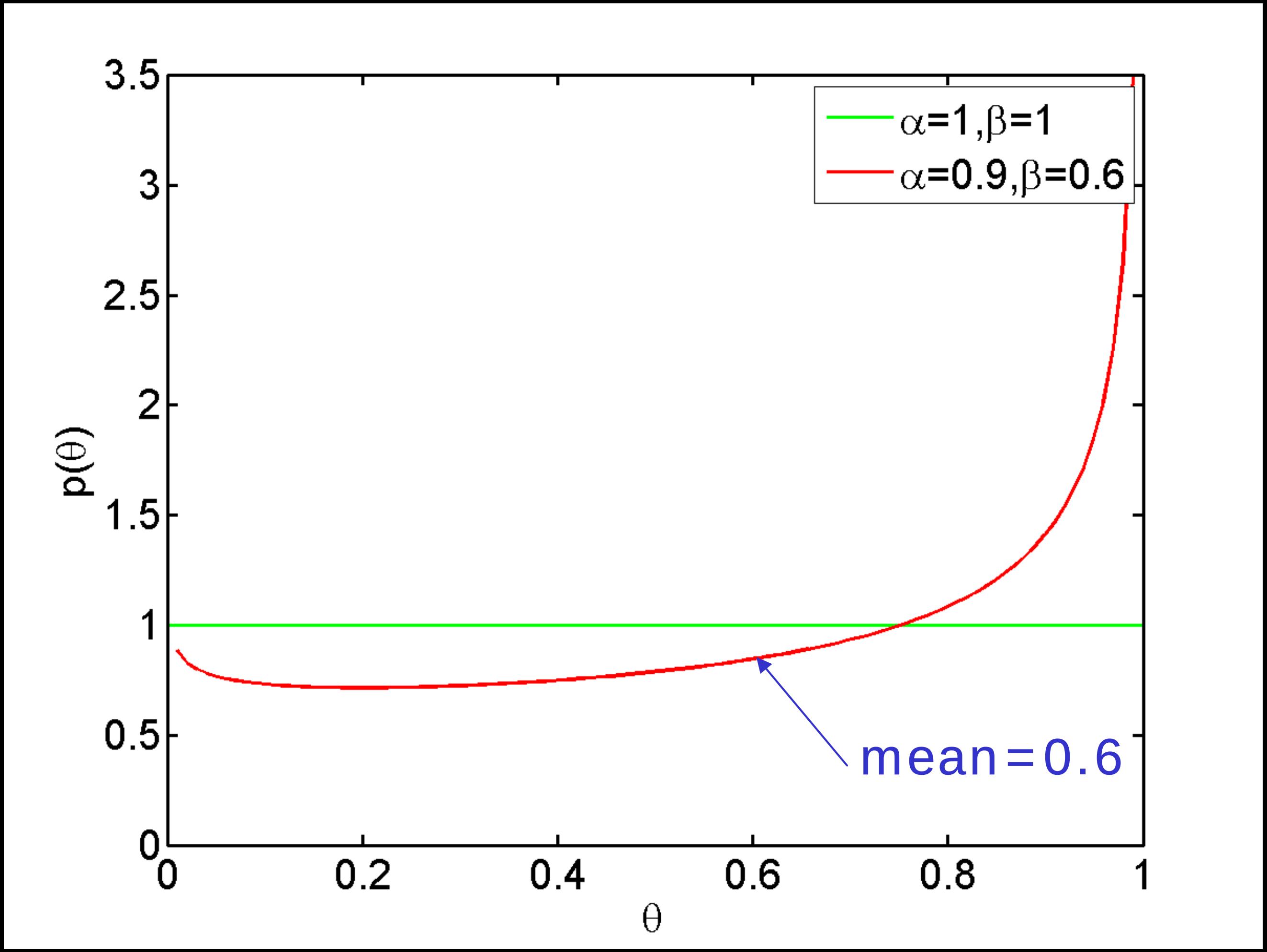
Likelihood

Prior

$$P(\theta \mid y) = \frac{P(y \mid \theta)p(\theta)}{P(y)}$$

Posterior

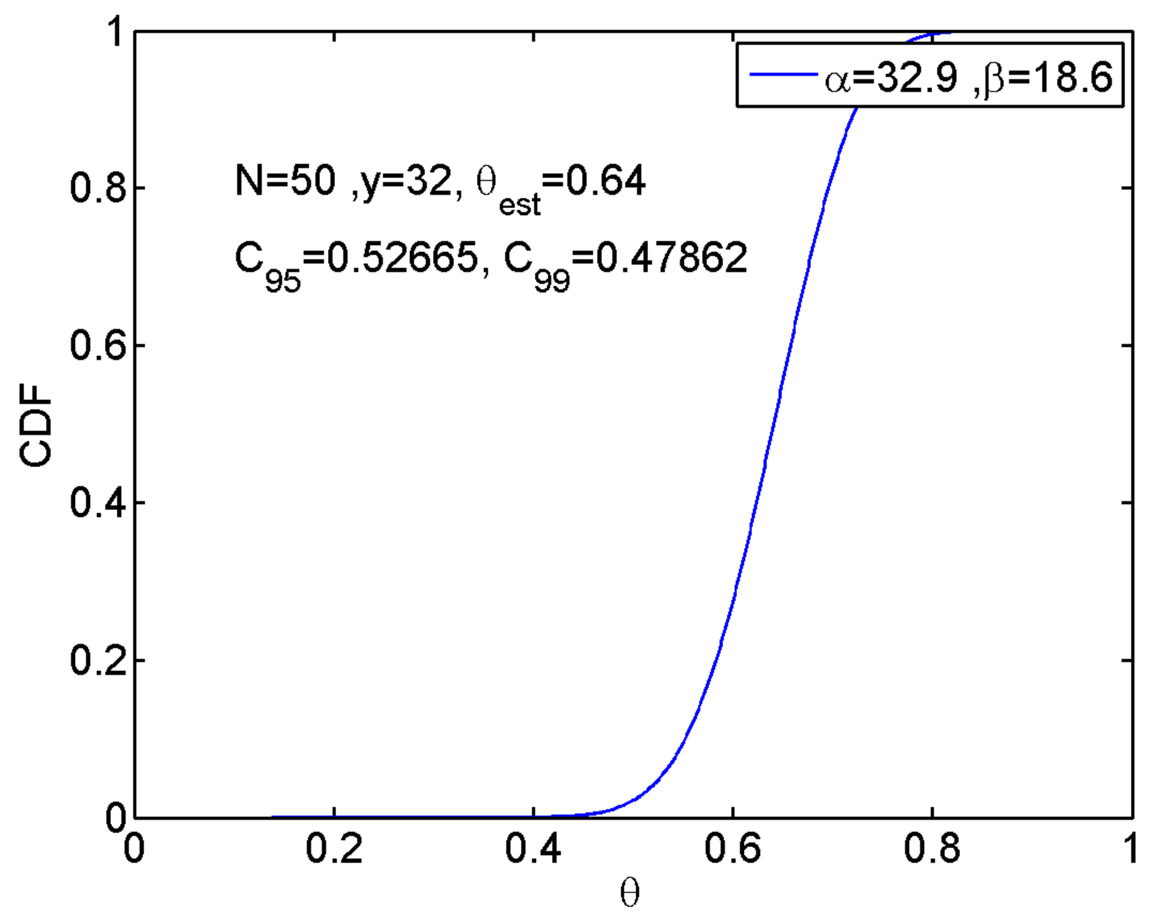
Prior distribution



mean=0.6

HPD credible sets – the Bayesian confidence interval

$$C_{1-\varepsilon} = \{\theta:\ P(\theta \mid y) \geq k(\varepsilon)\},\ \mathrm{CDF}(\theta \mid y) > 1-\varepsilon$$



The required number of samples *N*

- We need to be confident about the estimated detection probability

$$\text{Prob}(\theta > 99.6\%) = C_{1-\varepsilon}$$

	$C_{95\%}$	$C_{99\%}$
$\theta_{est} = 99.7\%$	9303	18994
$\theta_{est} = 99.8\%$	2285	3995

Uniform prior

	$C_{95\%}$	$C_{99\%}$
$\theta_{est} = 99.7\%$	8317	18301
$\theta_{est} = 99.8\%$	2147	3493

Informative prior

$\alpha=0.9$, $\beta=0.6$

Prior info reduces the need for samples

Credible sets when detecting 100%

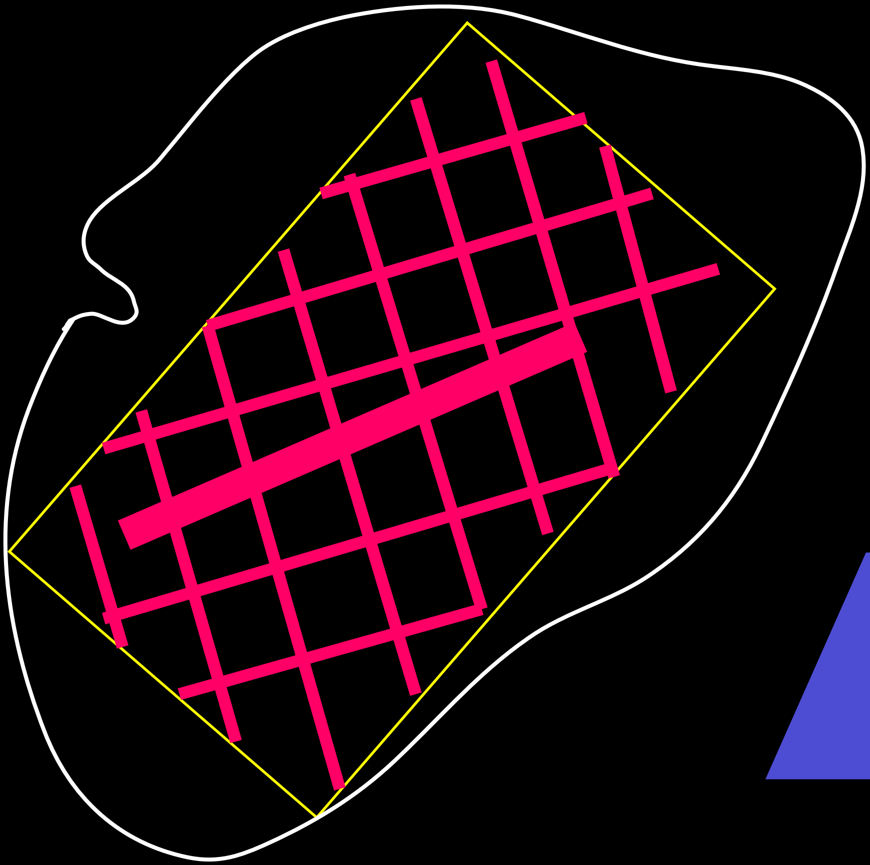
Minimum number of samples N

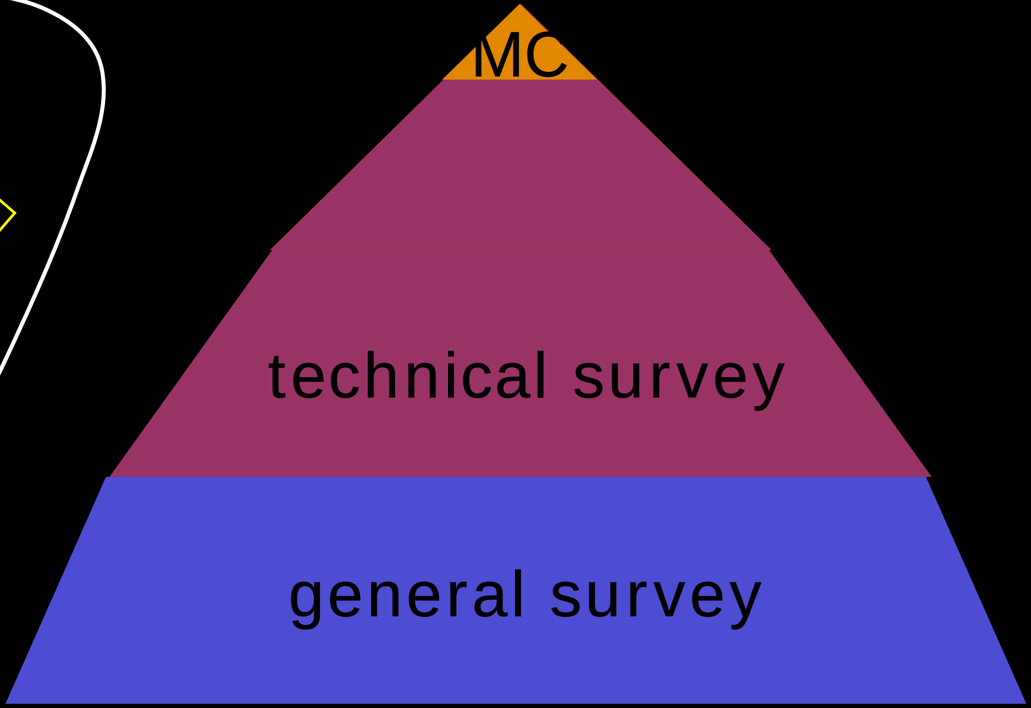
	Prob($\theta > 80\%$)	Prob($\theta > 99.6\%$)	Prob($\theta > 99.9\%$)
$C_{95\%}$	13	747	2994
$C_{99\%}$	20	1148	4602

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsense projects

Efficient MA by hierarchical approaches





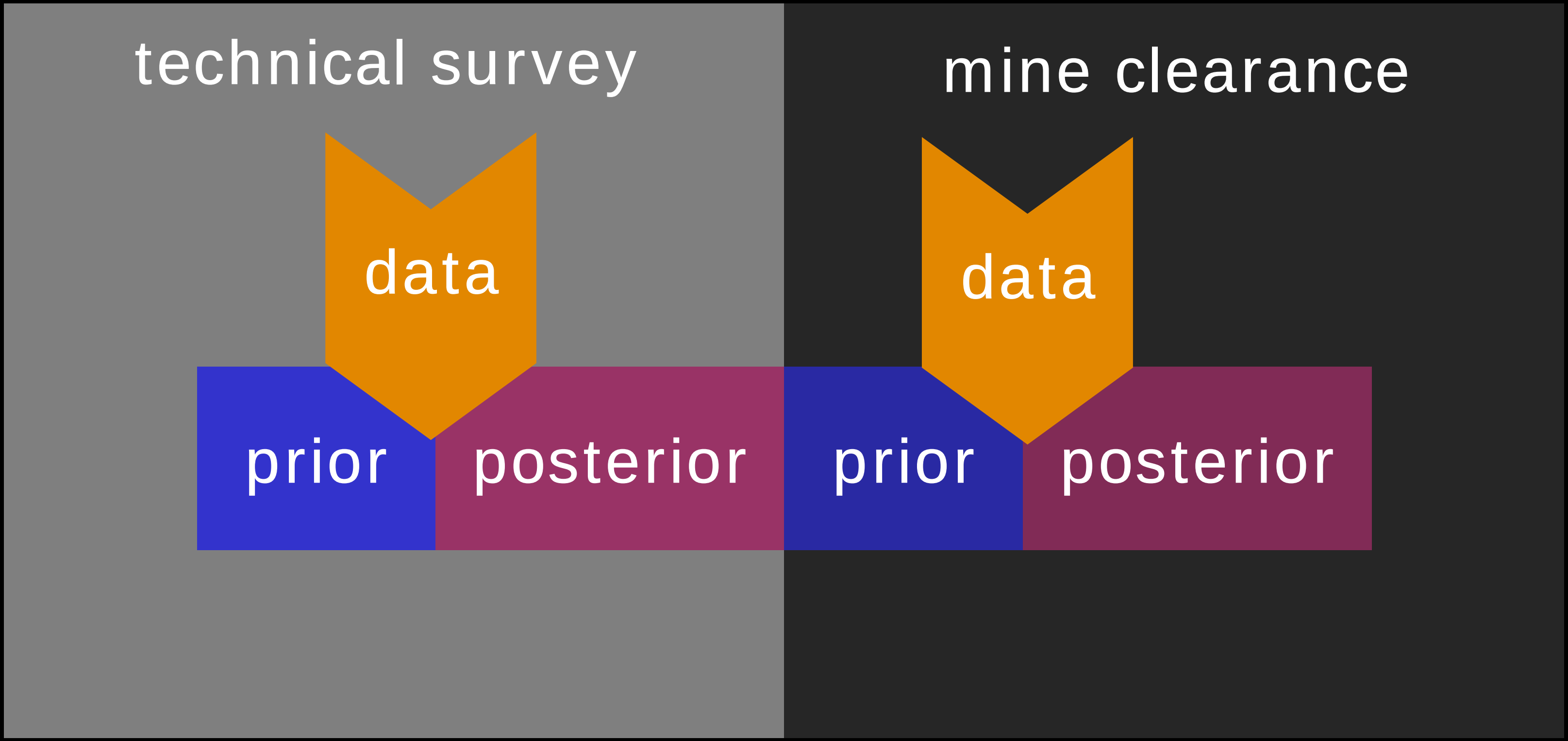
Refs: Håvard Bach, Paul Mackintosh

Danger maps

- The outcome of a hierarchical surveys
- Information about mine types, deployment patterns etc. should also be used
- Could be formulated/interpreted as a prior probability of mines

SMART system described in GICHD: Guidebook on Detection Technologies and Systems for Humanitarian Demining, 2006

Sequential information gathering



technical survey

data

prior

posterior

mine clearance

data

prior

posterior

Statistical information aggregation

- e=1 indicates encounter of a mine in a box at a specific location
- probability of encounter $P(e=1)$ from current danger map
- d=1 indicates detection by the detection system
- probability of detection $P(d=1)$ from current accreditation

$$P(e=1 \wedge d=0) = P(e=1)(1-P(d=1))$$

$$P(\text{no mine}) = 1 - P(e=1 \wedge d=0)$$

Statistical information aggregation

$$P(e=1)=0.2,\ \ P(d=1)=0.8$$
$$P(\text{no mine})=1-P(e=1\wedge d=0)=1-0.2*0.2=0.96$$

$$P(e=1)=1,\ \ P(d=1)=0.96$$
$$P(\text{no mine})=1-P(e=1\wedge d=0)=1-1*0.04=0.96$$

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsense projects

Where are we and how do we get further?

- No single existing method deliver sufficient detection performance
- No universal best method exists – every method has its pros and cons
- Fusion of sensors have been suggested in

  “Analysis and Fusion using Belief Function Theory of Multisensor Data for Close-range Humanitarian Mine Detection.

  PhD Thesis RMA, 2001 by Nada Milisavljević

Does not immediately apply to fusion of heterogenous methods

Advantages

- Combination leads to a possible exponential increase in detection performance
- Combination leads to better robustness against changes in environmental conditions

Challenges

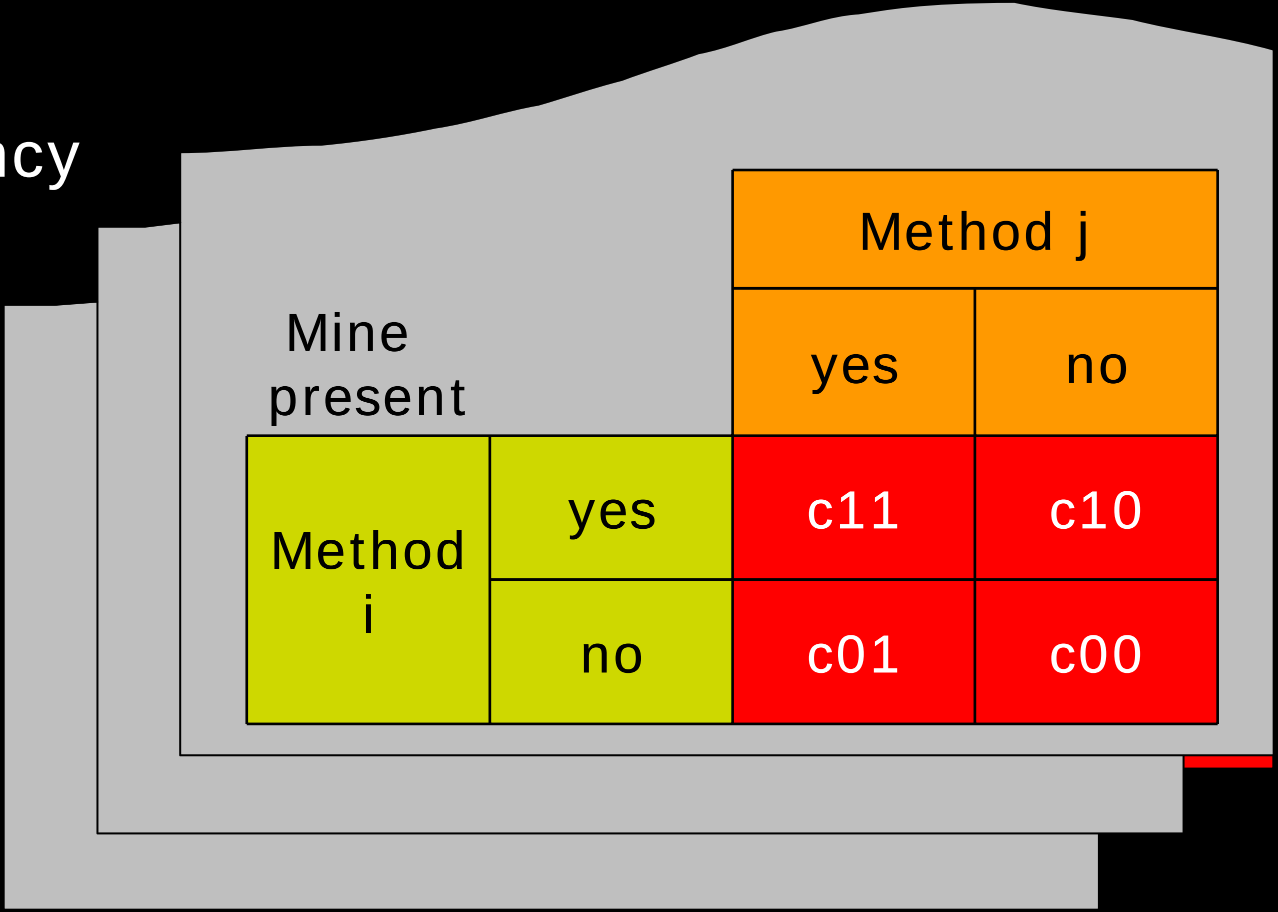
- Need for certification procedure of equipment under well-specified conditions (ala ISO)
- Need for new procedures which estimate statistical dependences between existing methods
- Need for new procedures for statistically optimal combination

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsense projects

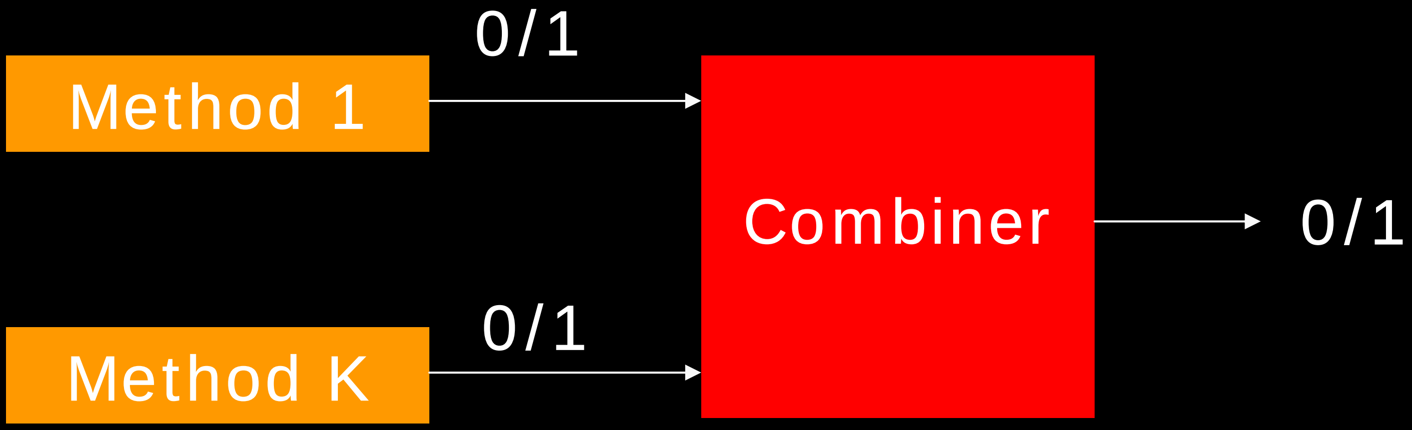
Dependencies between methods

Contingency tables



Mine present		Method j	
		yes	no
Method i	yes	c11	c10
	no	c01	c00

Optimal combination



Optimal combination depends on contingency tables

Optimal combiner

Method		Combiner						
1	**2**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	0						0	0
0	1	0	0	0	1		1	1
1	0	0	1	1	0	0	1	1
1	1	1	0	1	0	1	0	1

OR rule is optimal for independent methods

$2^{2^{K-1}} - 1$ possible combiners

Example

$$p_{d1} = 0.8, p_{fa1} = 0.1 \quad p_{d2} = 0.7, p_{fa2} = 0.1$$

$$p_d = 1 - (1 - 0.8) \cdot (1 - 0.7) = 0.94$$

$$p_{fa} = 1 - (1 - 0.1) \cdot (1 - 0.1) = 0.19$$

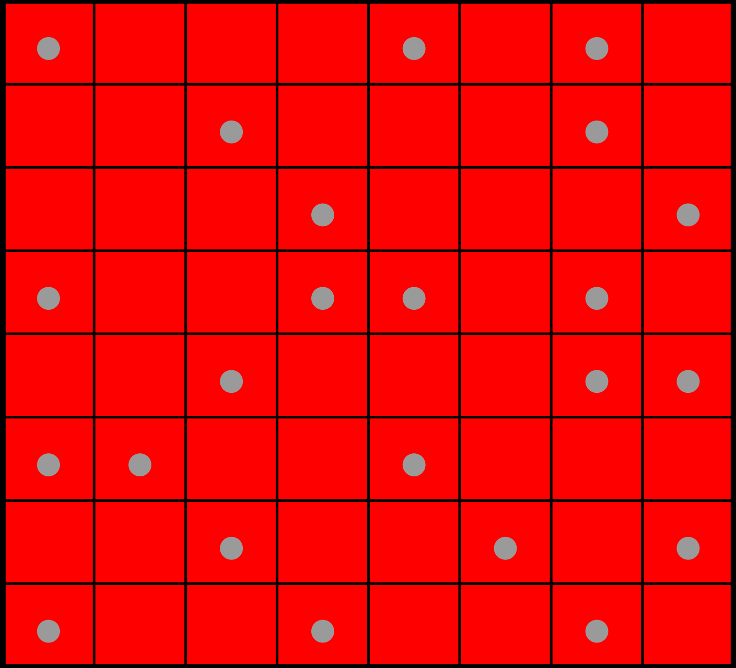
Exponential increase in detection rate

Linear increase in false alarm rate

Joint discussions with: Bjarne Haugstad

Artificial example

- N=23 mines
- Method 1 (flail): P(detection)=0.8, P(false alarm)=0.1
- Method 2 (metal detector): P(detection)=0.7, P(false alarm)=0.1
- Resolution: 64 cells



How does detection and false alarm rate influence the possibility of gaining by combining methods?

Confusion matrix for method 1

		True	
		yes	no
Estimated	yes	19	5
	no	4	36

Confidence of estimated detection rate

- With *N*=23 mines 95%-credible intervals for detection rates are extremely large!!!!

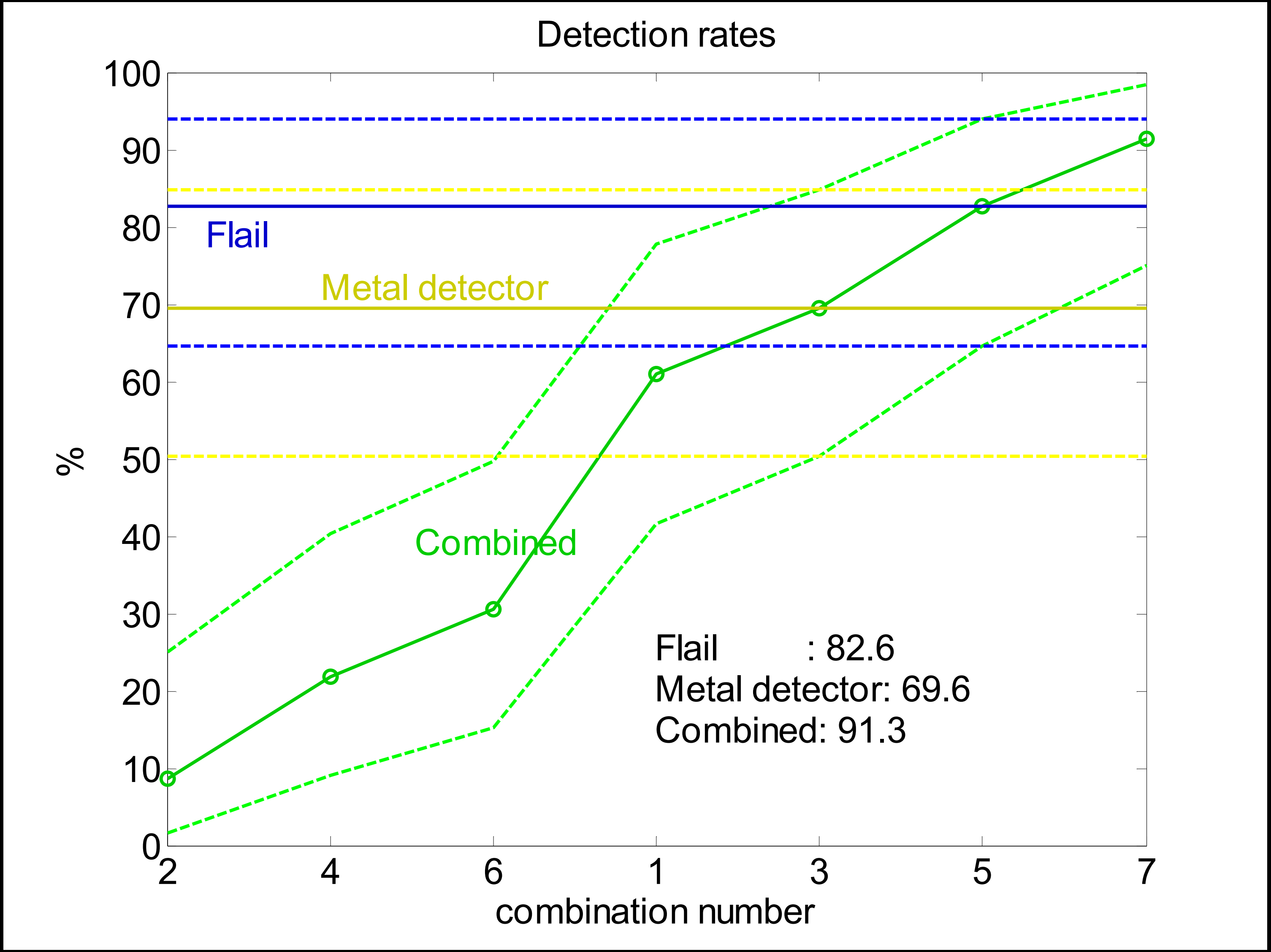
$$[64.5\% \quad 82.6\% \quad 93.8\%]$$

$$[50.4\% \quad 69.6\% \quad 84.8\%]$$

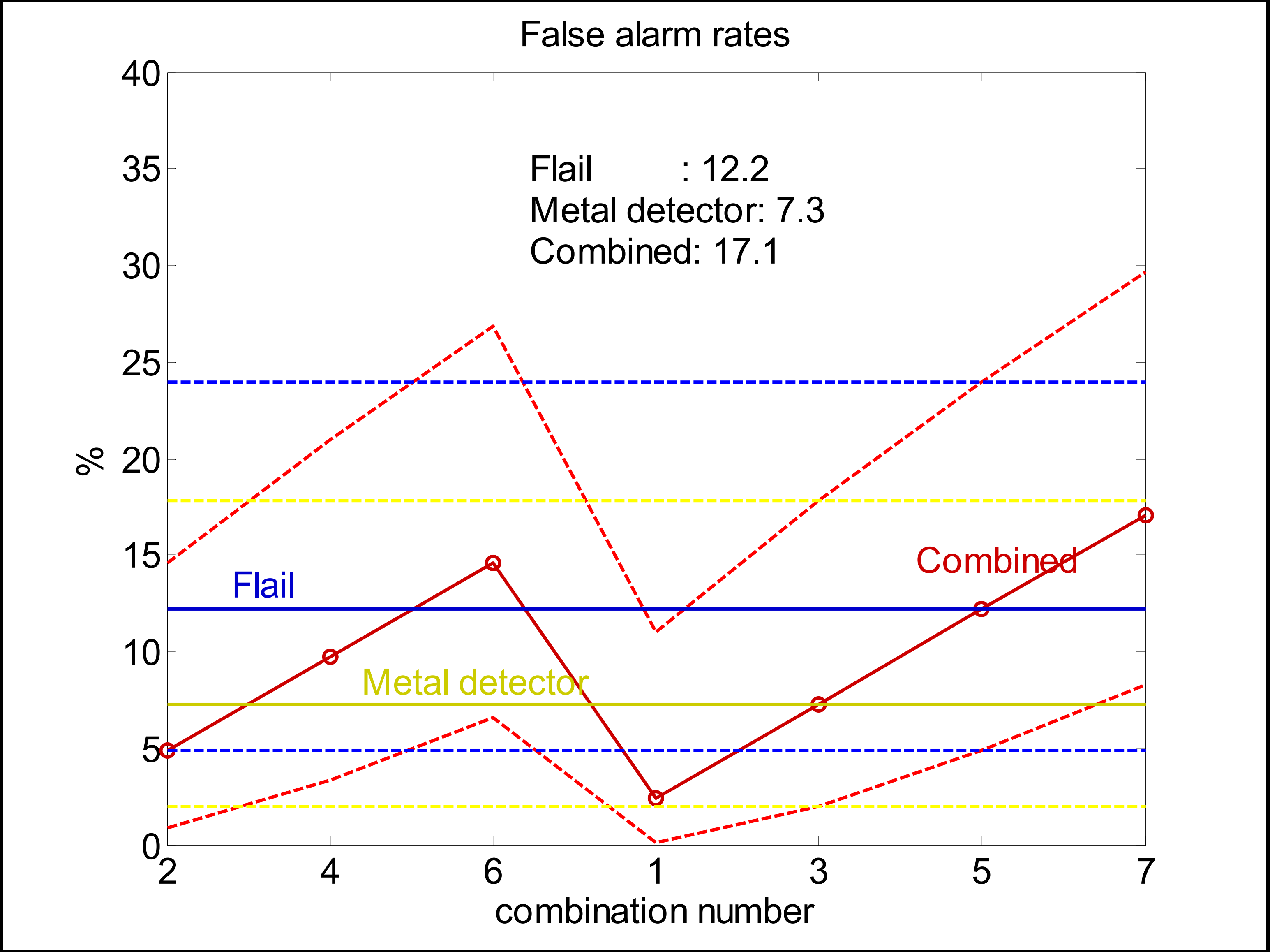
Confidence for false alarm rates

- Determined by deployed resolution
- Large resolution - many cells gives many possibilities to evaluate false alarm.
- In present case: 64-23=41 non-mine cells

$$[4.9\% \quad 12.2\% \quad 24.0\%]$$



Flail : 82.6
Metal detector: 69.6
Combined: 91.3


False alarm rates
Flail : 12.2
Metal detector: 7.3
Combined: 17.1
%
0
5
10
15
20
25
30
35
40
Flail
Metal detector
Combined
2
4
6
1
3
5
7
combination number

Comparing methods

- Is the combined method better than any of the two orginal?
- Since methods are evaluated on same data a paired statistical McNemar with improved power is useful

Method1 (flail): 82.6% < 91.3% Combined 🚫

Method2 (MD): 69.6% < 91.3% Combined ☺

Outline

- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination of methods
  - Advantages
  - Methodology
  - DeFuse and Xsens projects

DeFuse scientific objectives

- Obtain general scientific knowledge about the advantages of deploying a combined approach
- Eliminate confounding factors through careful experimental design and specific scientific hypotheses
- Test the general scientific hypothesis is that there is little dependence between missed detections in successive runs of the same or different methods
- To accept the hypothesis under varying detection/clearance probability levels
- To lay the foundation for new practices for mine action, but it is not within scope of the pilot project

Xsense

- The scope of the Xsense program is to realize a reliable, sensitive, portable and low-cost explosive detector
- The detector will be miniaturized and will therefore be highly suitable for use in anti terror efforts, boarder control, environmental monitoring and demining
- The sensitivity will be optimized by a concentrated effort in data processing (reducing noise and pattern recognition) and emerging sensing principles
- The reliability of the detector will be ensured by combining several independent sensor technologies

Conclusions

- A cross-disciplinary effort is required to obtain sufficient knowledge about physical, operational and processing possibilities and constraints as well as clear definition of a measurable goal – the right tool for the right problem
- Statistical decision theory and modeling is essential for optimal use of prior information and empirical evidence
- It is very hard to assess the necessary high performance which is required to have a tolerable risk of casualty
- The use of sequential information aggregation is promising for developing new hierarchical survey schemes (SOPs)
- Combination of methods is a promising avenue to overcome curr