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

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

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
- Statistic modeling usuallay focuses on identifying important parameters
- machine learning learns complex models from collections of data to make optimal predictions in new situations

Why do we need statistical models and machine learning?

- statistical modeling is the principled framework to handle uncertainty and complexity
- Statistic modeling usually focuses on identifying important parameters
- machine learning learns complex models from collections of data to make optimal predictions in new situations

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 statistical methods sometimes wrongly leads to the conclusion that they are of little practical use

After the dogs went in we never saw an accident

Most suspected areas have very few mines

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

- Pitfalls and misuse of statistical methods sometimes wrongly leads to the conclusion that they are of little pra[illegible]

Some data are in the tail of the distribution: generalization from few examples is not possible

Smoking is not dangerous: my granny just turned 95 and has been a heavy smoker all his live

The elements of statistical decision theory

Data
- Sensor
- Calibration
- Post clearance
- External factors

Prior knowledge
- Physical knowledge
- Experience
- Environment

Statistical models

Loss 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 project

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

$$Frequency = \frac{996}{1000} = 99,6\%$$

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

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

Detection probability - tossing a coin

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

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

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

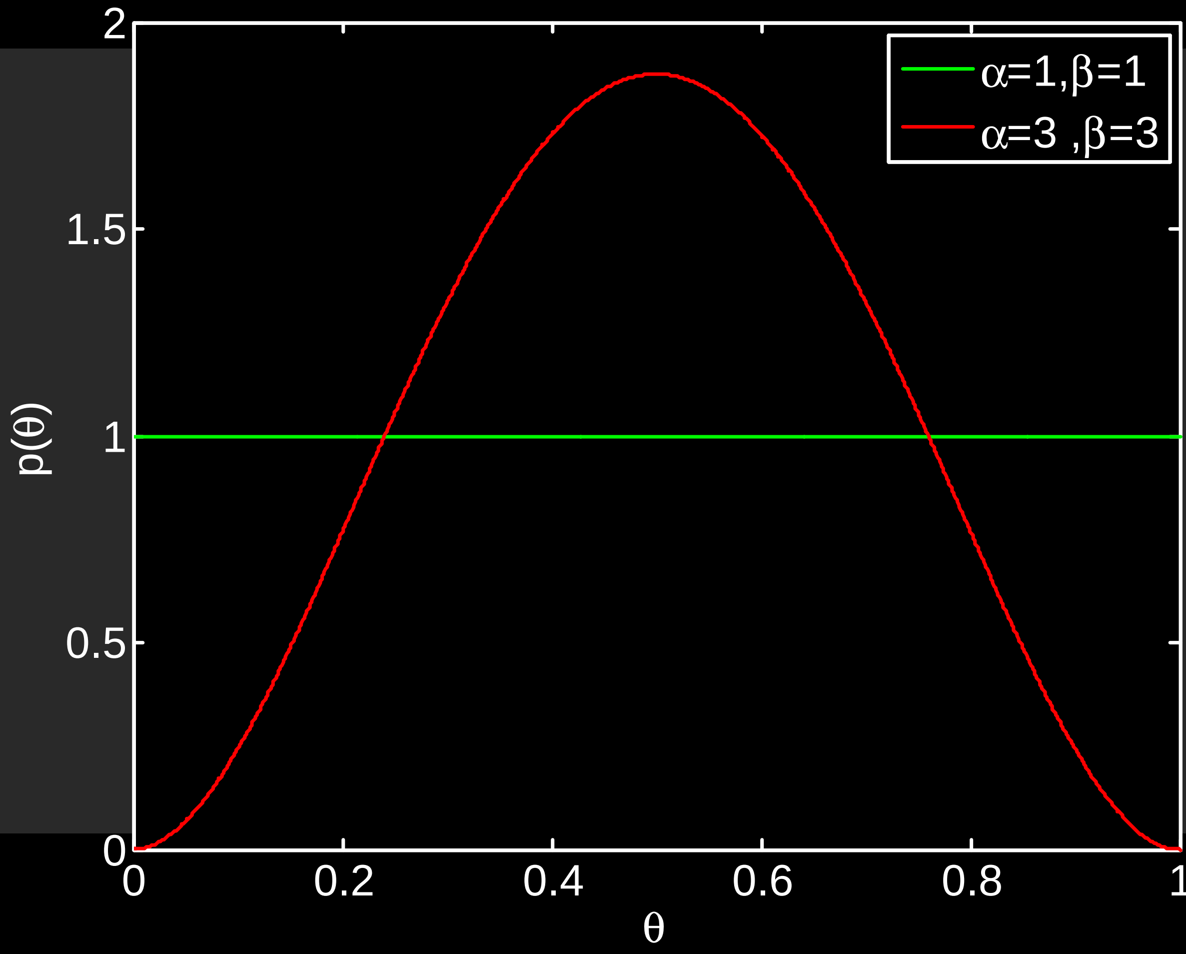
Data likelihood

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

Likelihood

Prior

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

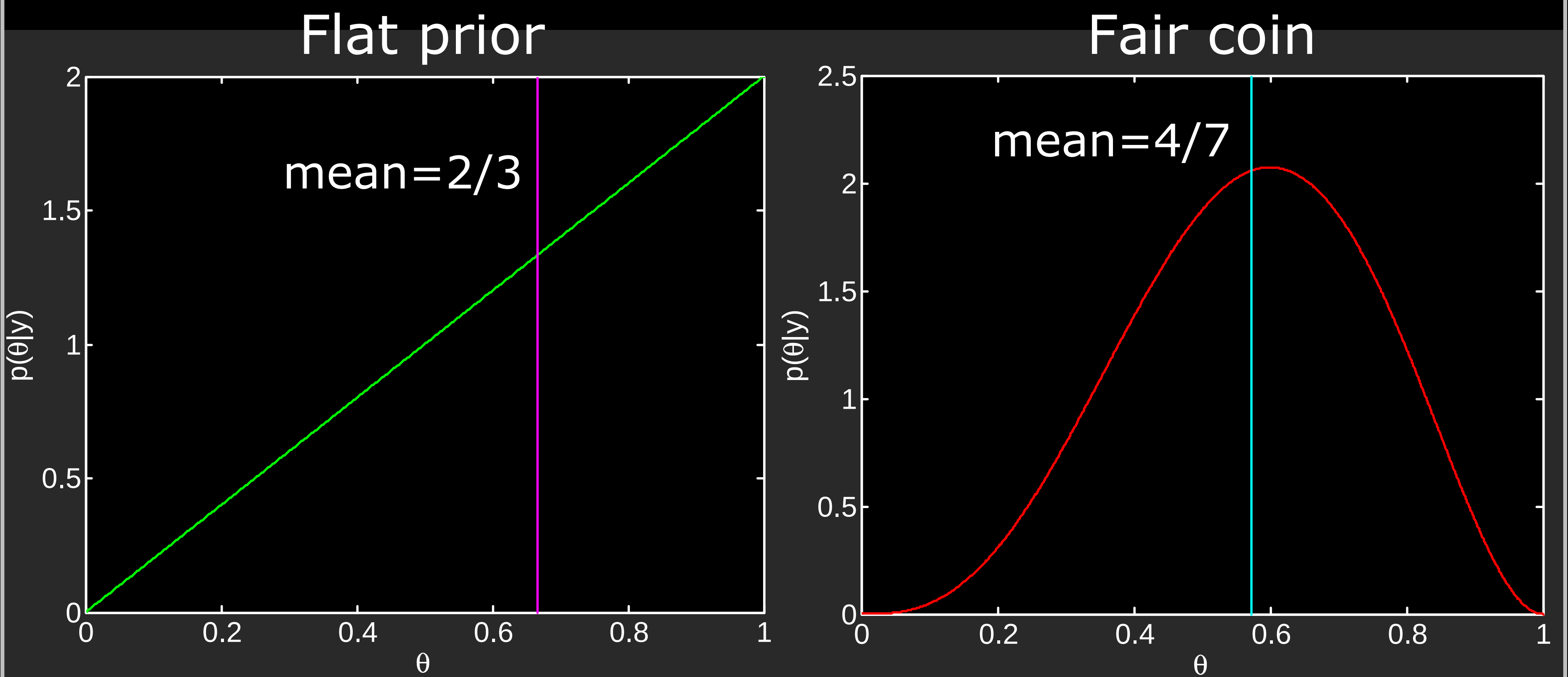
Posterior

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

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



$Beta(\theta \mid 2,1)$ $Beta(\theta \mid 4,3)$

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 project

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 analysis – e.g. some areas might better be fenced than cleared
- Need for professional risk analysis, communication management and control involving all partners (MAC, NGOs, commercial etc.)

What are the requirements for mine action risk

- Tolerable risk for individuals comparable to other natural risks
-
- management and ... C, NGOs, commercia...

Fact

99.6% 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 behavior pattern provides the probability of encounter: Prob(mine encounter)

Prob(casualty)=(1-Prob(detect)) * Prob (mine encounter)

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

Requirements on detection probability

Prob(causality)=(1-Prob(detection))*Prob(encounter)

Prob(detection)=1-Prob(causality)/Prob(encounter)

- Prob(encounter)= ρ*a
  - ρ : homogeneous mine density (mines/m^2), a: yearly footprint area (m^2)
- Prob(causality)=10^{-5} per year

Maximum yearly footprint area in m^2

P(detection)	ρ : mine density (mines/km²)				
	0.1	1	10	100	1000
0.996	25000	2500	250	25	2.5
0.9	1000	100	10	1	0.1

Reference: Bjarne Haugstad, FFI

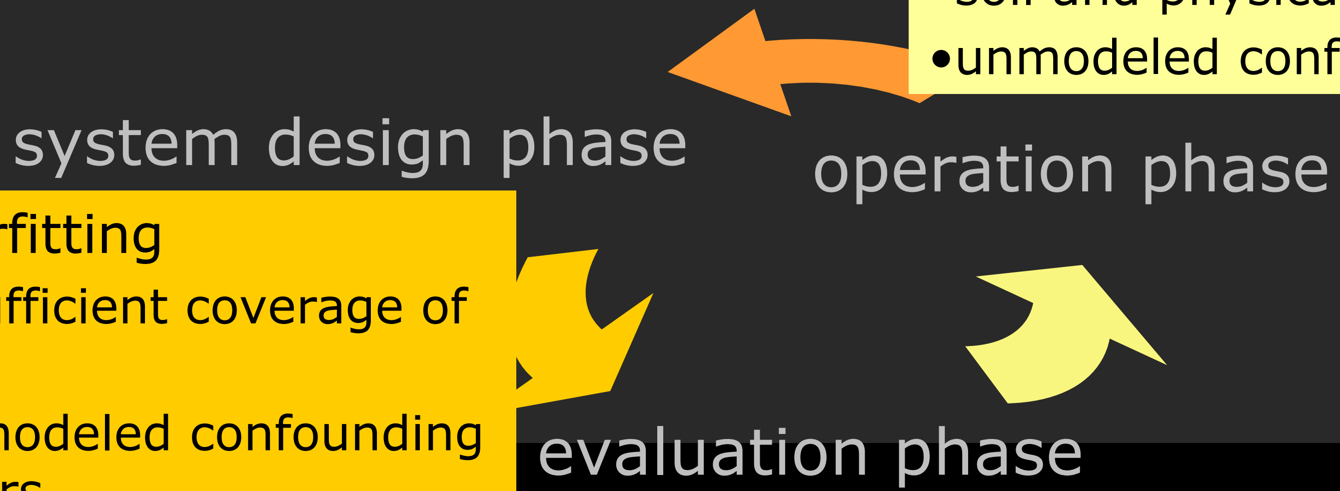
Outline

- 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 project

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



Overfitting
•insufficient coverage of data
•unmodeled confounding factors
•unsufficient model fusion and selection

Two types of error in detection of mines

Sensing error

The system does not sense the presence of the mine object

decrease in detection probability

Decision error

The detector misinterprets the sensed signal

increase in false alarm rate

Two types of error in detection of mines

Example: metal detector

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

Example: mine detection dog

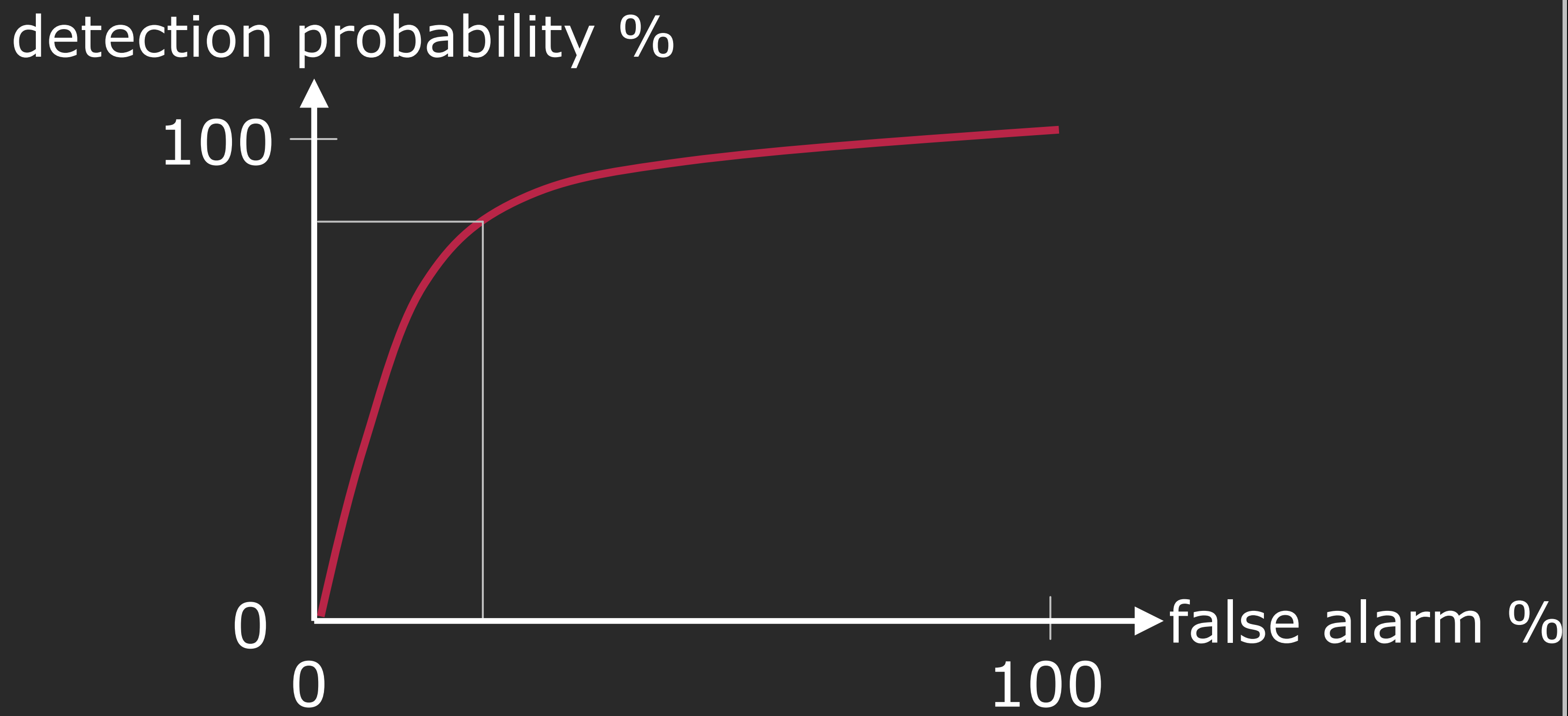
- Sensing error: the TNT leakage from the mine was too low
- Decision error: the dog handler misinterpreted the dogs indication

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)



Inferring the detection probability

- N independent mine areas for evaluation
- y detections observed
- true detection probability θ

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

Bayes rule: combining data likelihood and prior

Likelihood

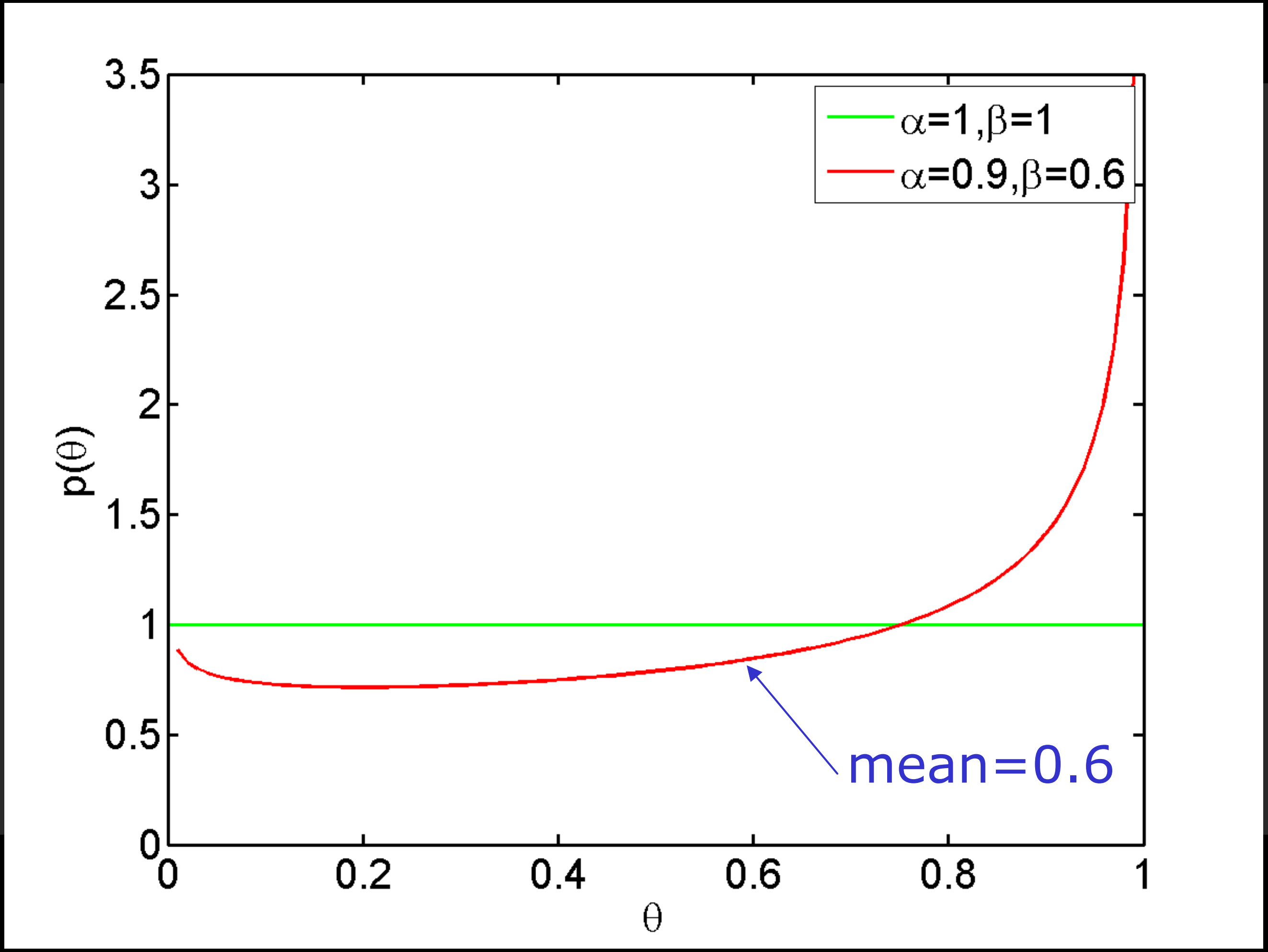
Prior

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

Posterior

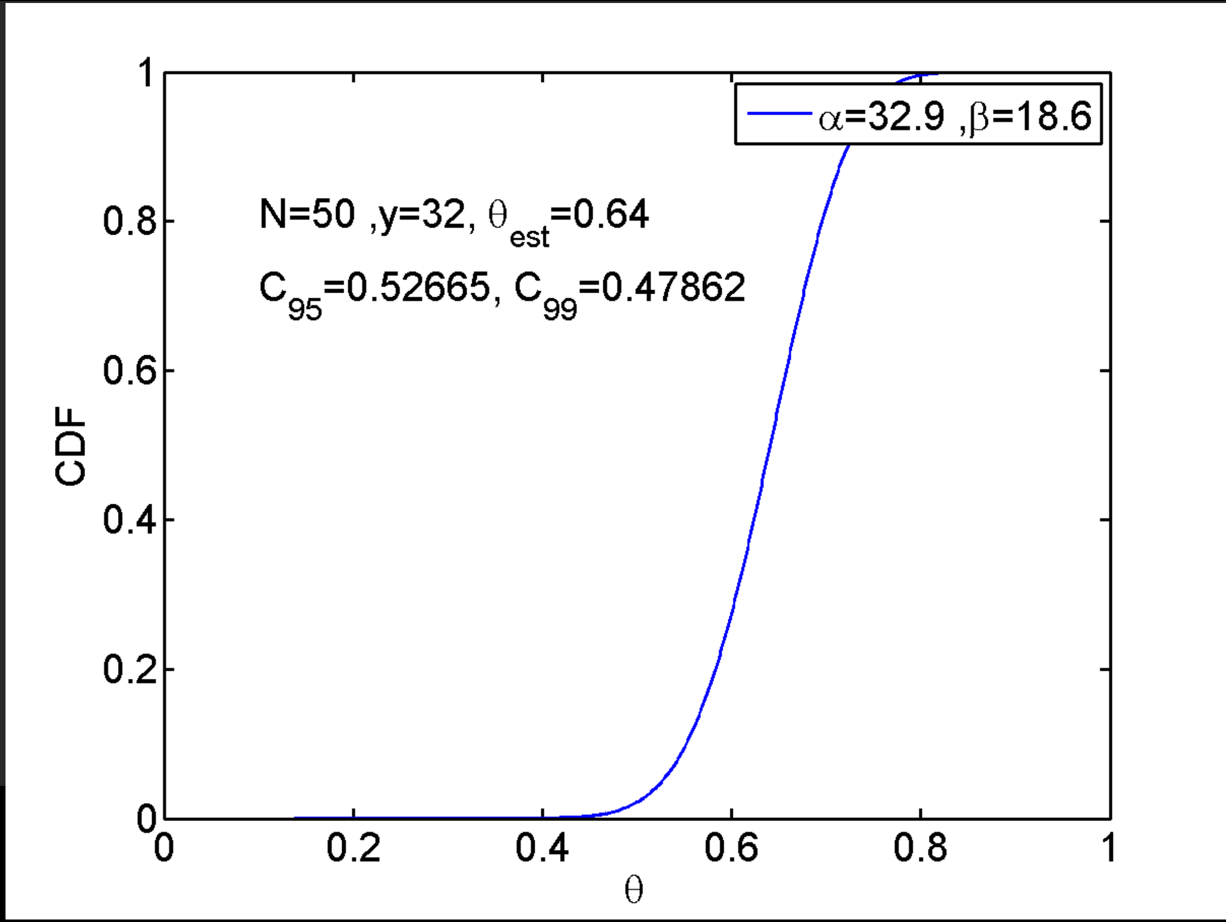
$$P(\theta \mid y) = Beta(\theta \mid y+\alpha, \beta+n-y) \sim \theta^{y+\alpha}\theta^{n-y+\beta}$$

Prior distribution



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

α=0.9, β=0.6

Credible sets when detecting 100%

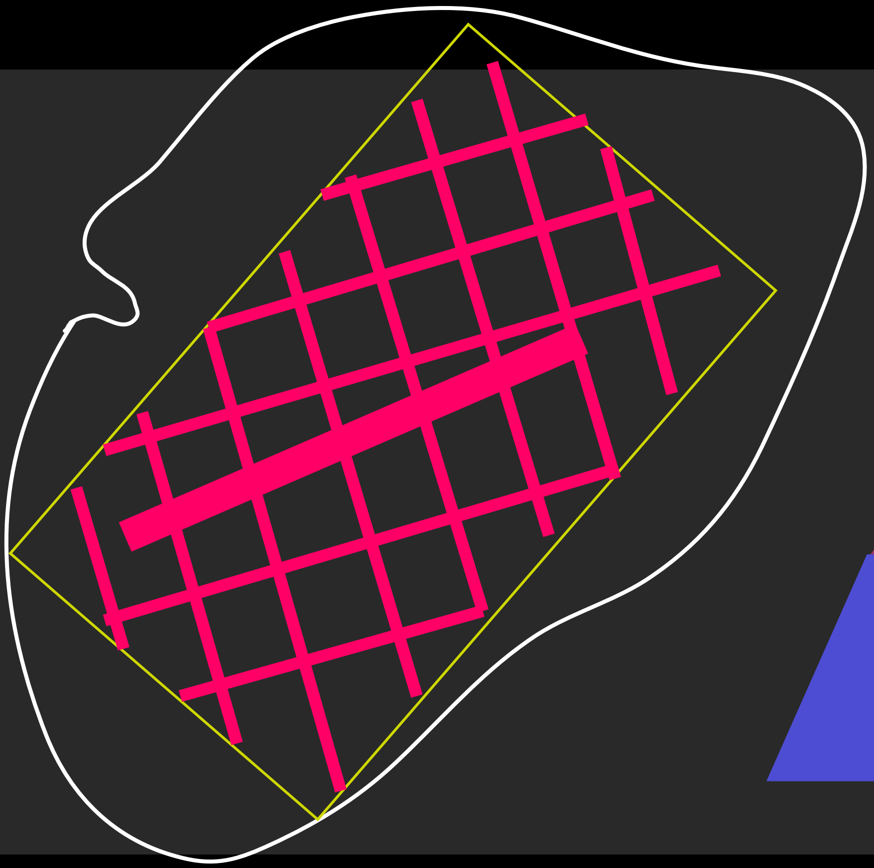
Minimum number of samples N

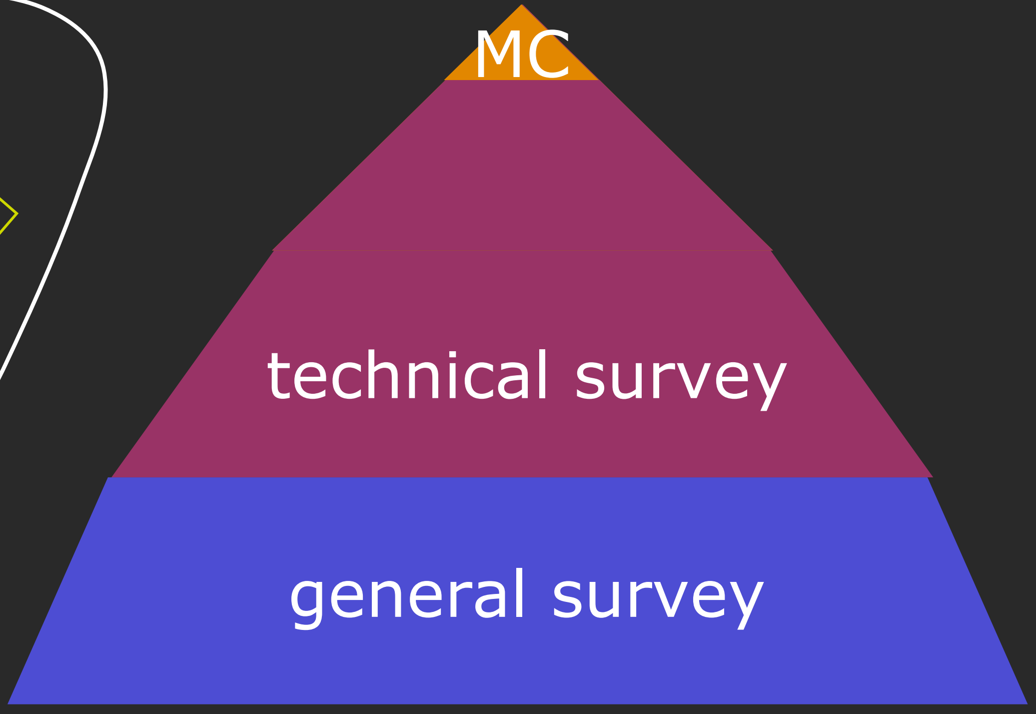
	Prob($\theta > 80\%$)	Prob($\theta > 99.6\%$)	Prob($\theta > 99.9\%$)
$C_{95\%}$	13	747	2994
$C_{99\%}$	20	1148	4602

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 project

Efficient MA by hierarchical approaches





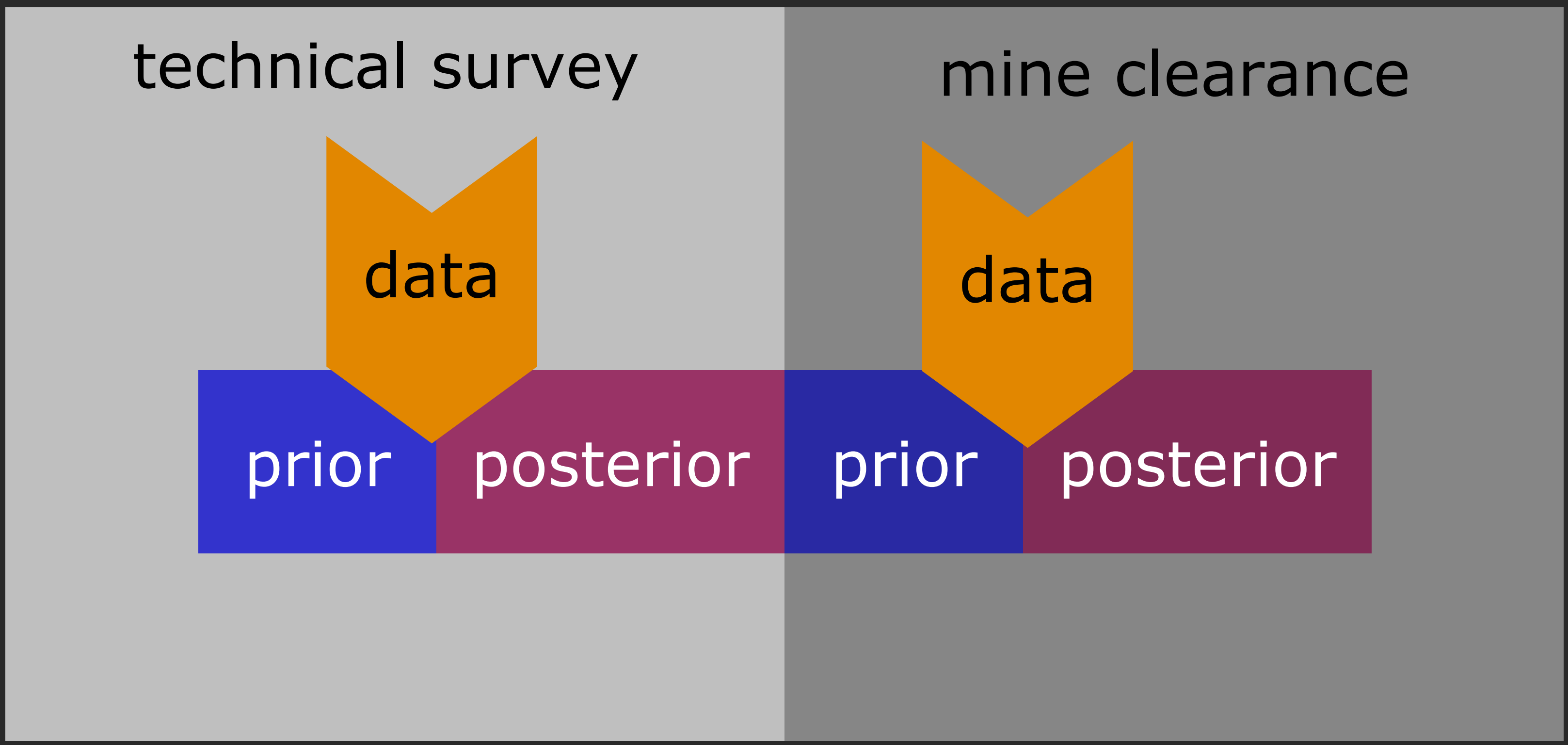
Ref: Håvard Bach, Paul Mackintosh

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



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

Example: flail in a low danger area

$$P(e=1)=0.2,\ P(d=1)=0.8$$

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

Example: manual raking in a high danger area

$$P(e=1)=1,\ P(d=1)=0.96$$

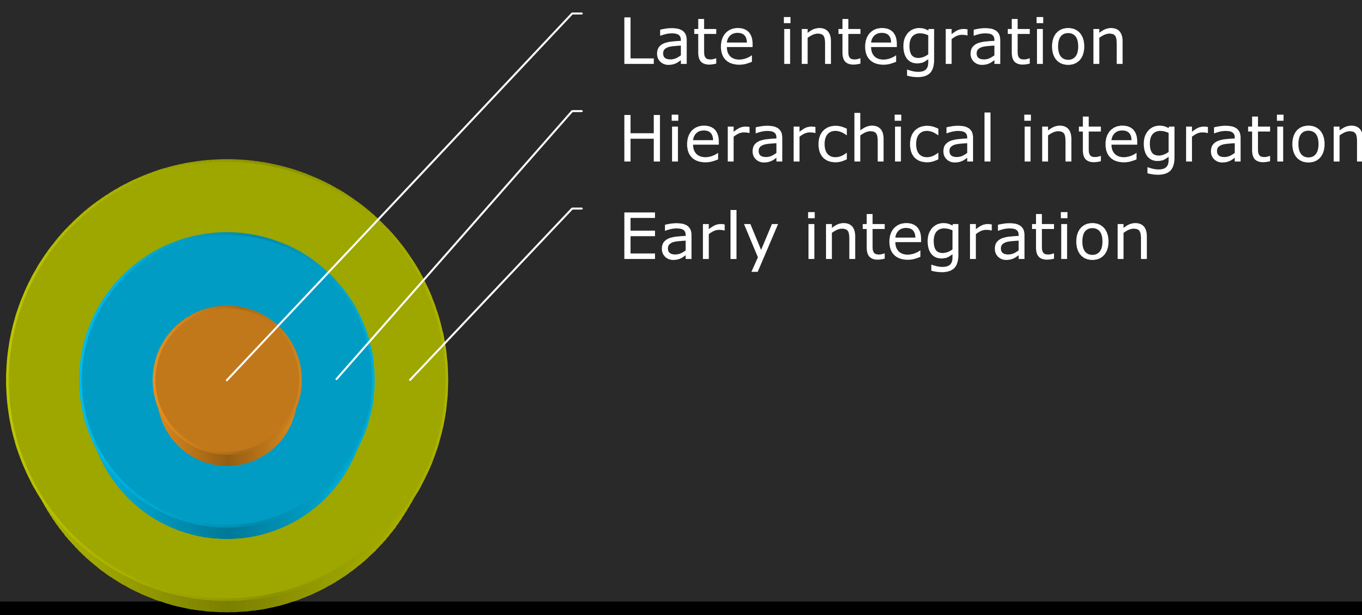
$$P(\text{no mine})=1-P(e=1 \wedge d=0)=1-1*0.04=0.96$$

Outline

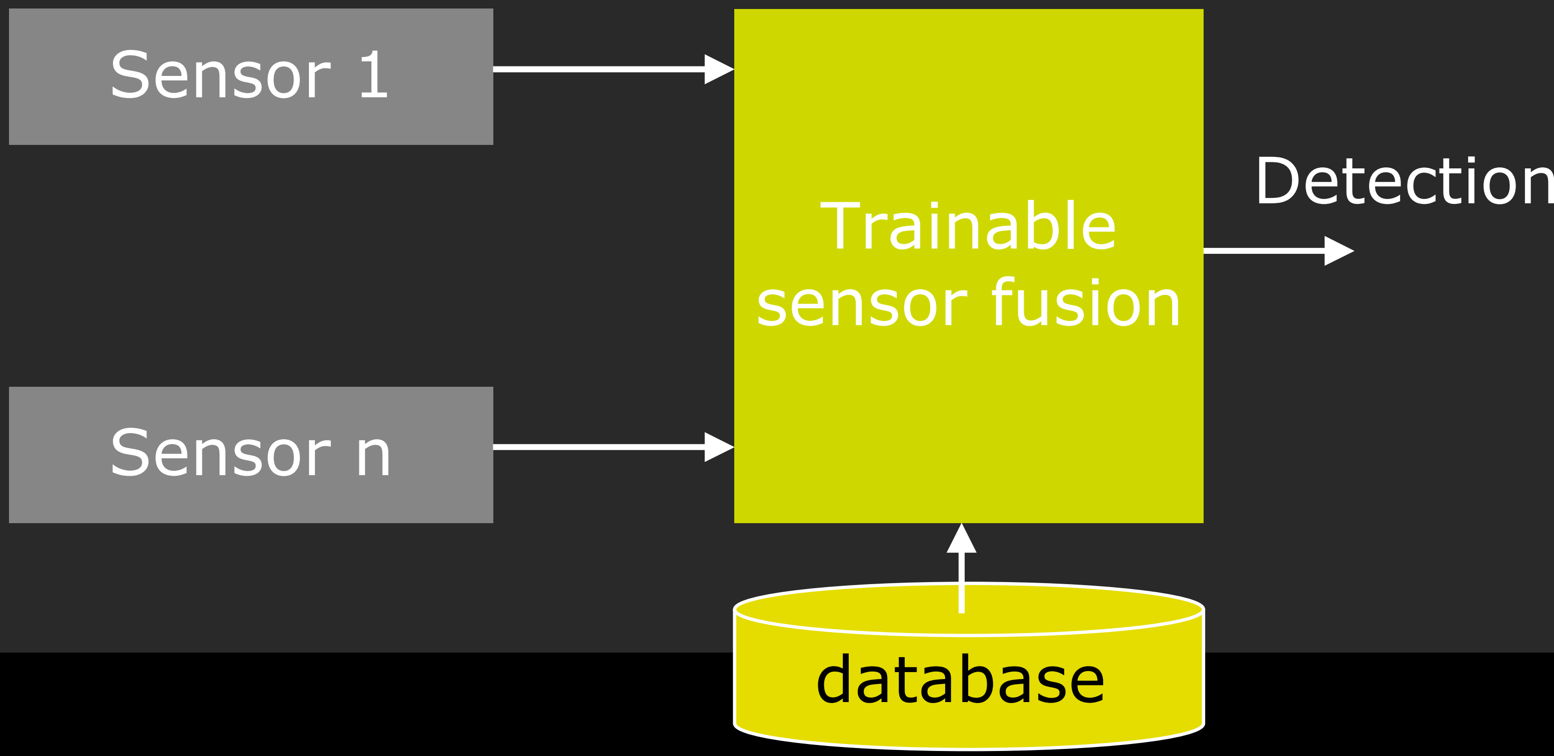
- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination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

Improving performance by fusion of methods

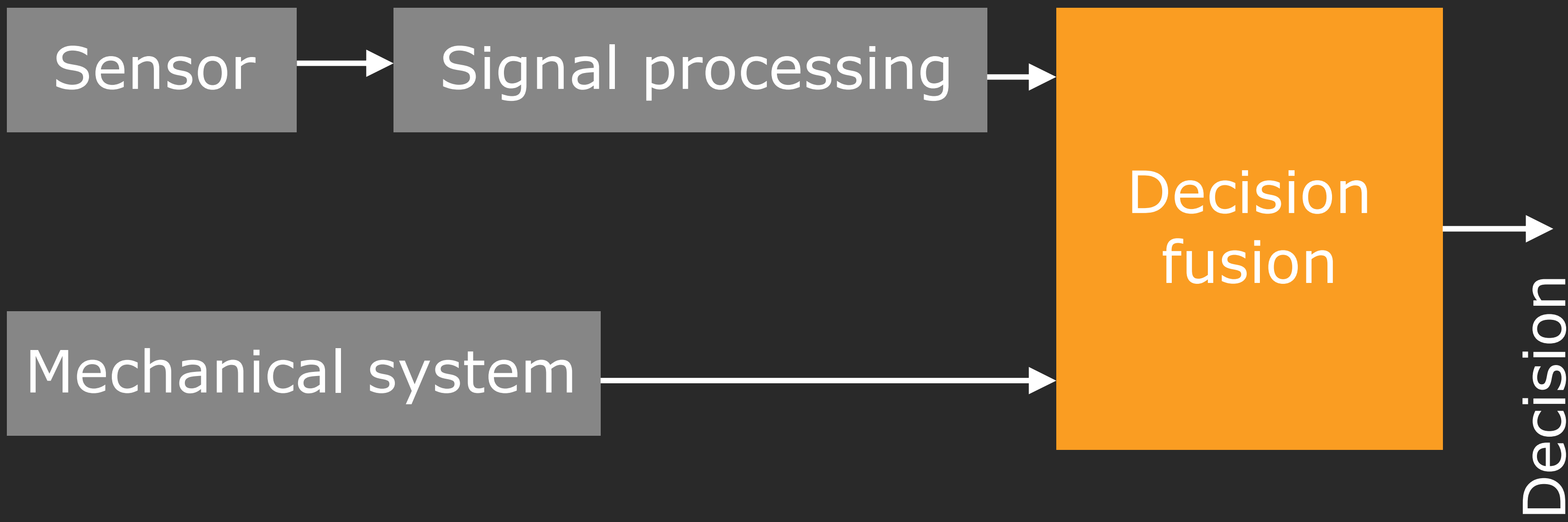
- Methods (sensors, mechanical etc.) supplement each other by exploiting different aspect of physical environment



Early integration – sensor fusion



Late integration – decision fusion



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

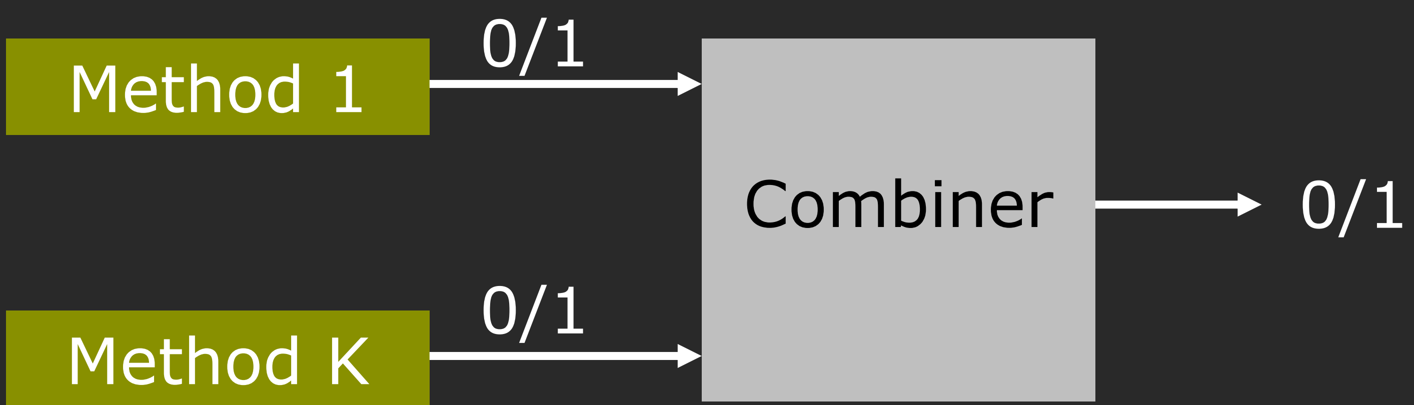
- The design and evaluation of mine clearance equipment – the problem of reliability
  - Detection probability – tossing a coin
  - Requirements in mine action
  - Detection probability and confidence in MA
  - Using statistics in area reduction
- Improving performance by information fusion and combination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

Dependencies between methods

Contingency tables

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

Optimal combination



Optimal combiner depends on contingency tables

Optimal combiner

OR rule is optimal for independent methods

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

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

OR rule is optimal for independent methods

Method 1:	1	0	0	1	0	0	1	0	1	0
Method 2:	0	1	0	0	1	0	1	1	1	0
Combined:	1	1	0	1	1	0	1	1	1	0

$$P_d(OR) = P(\hat{y}_1 \vee \hat{y}_2 = 1 \mid y = 1)$$
$$= 1 - P(\hat{y}_1 = 0 \wedge \hat{y}_2 = 0 \mid y = 1)$$
$$= 1 - P(\hat{y}_1 = 0 \mid y = 1) \cdot P(\hat{y}_2 = 0 \mid y = 1)$$
$$= 1 - (1 - P_{d1}) \cdot (1 - P_{d2})$$

independence

False alarm follows a similar rule

$$
\begin{aligned}
&P_{fa}(OR) = \\
&P(\hat{y}_1 \vee \hat{y}_2 = 1 \mid y = 0) \\
&= 1 - P(\hat{y}_1 = 0 \wedge \hat{y}_2 = 0 \mid y = 0) \\
&= 1 - P(\hat{y}_1 = 0 \mid y = 0) \cdot P(\hat{y}_2 = 0 \mid y = 0) \\
&= 1 - (1 - P_{fa1}) \cdot (1 - P_{fa2})
\end{aligned}
$$

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

Testing independence – Fisher's exact test

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

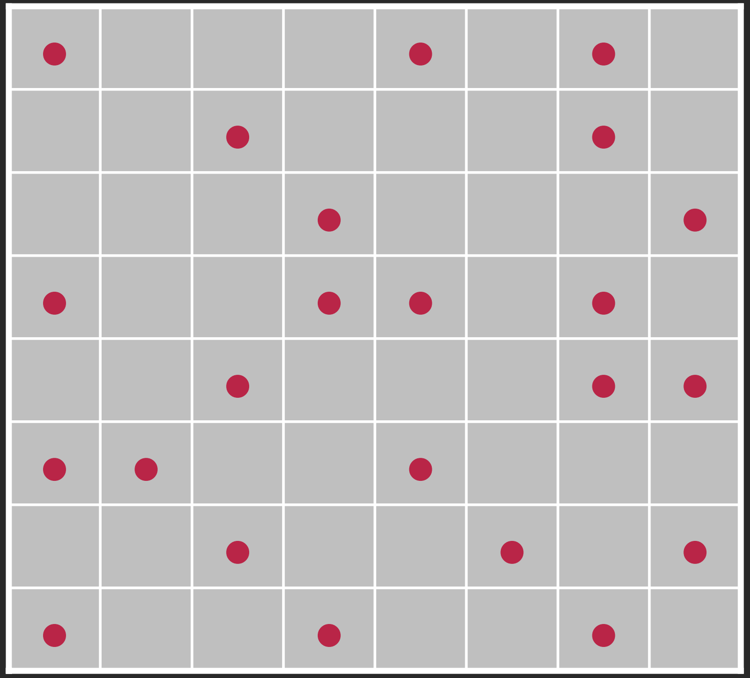
- Hypothesis: Method i and j are independent
- Alternatives: Dependent or positively (negatively) correlated

$$\mathrm{H}: P(\hat{y}_i = 0, \hat{y}_j = 0) = P(\hat{y}_i = 0) \cdot P(\hat{y}_j = 0)$$

$$\mathrm{A}: P(\hat{y}_i = 0, \hat{y}_j = 0) > P(\hat{y}_i = 0) \cdot P(\hat{y}_j = 0)$$

Artificial example

- N=23 mines
- Method 1: P(detection)=0.8, P(false alarm)=0.1
- Method 2: 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!!!!

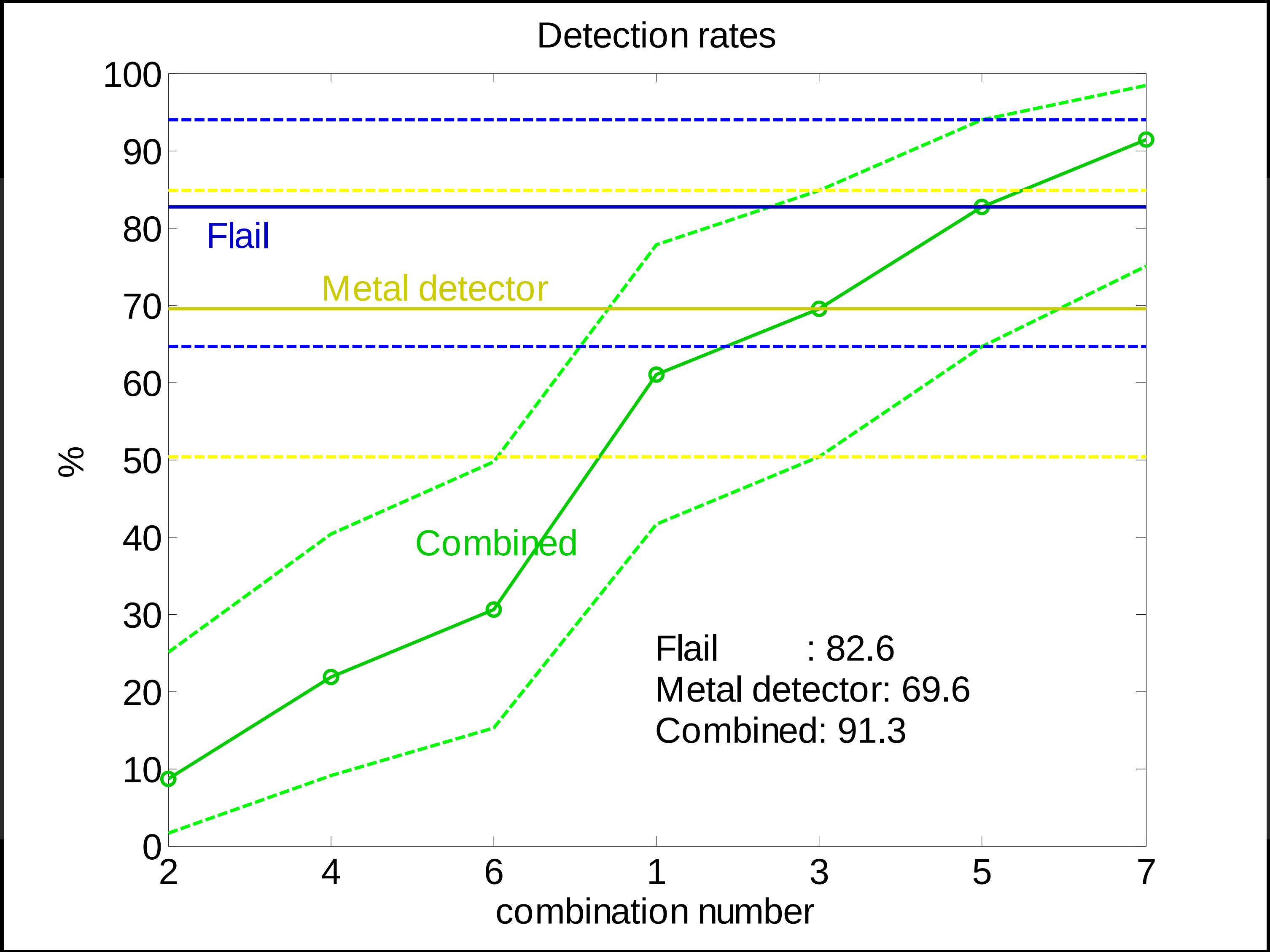
Method1 (flail): [64.5% 82.6% 93.8%]

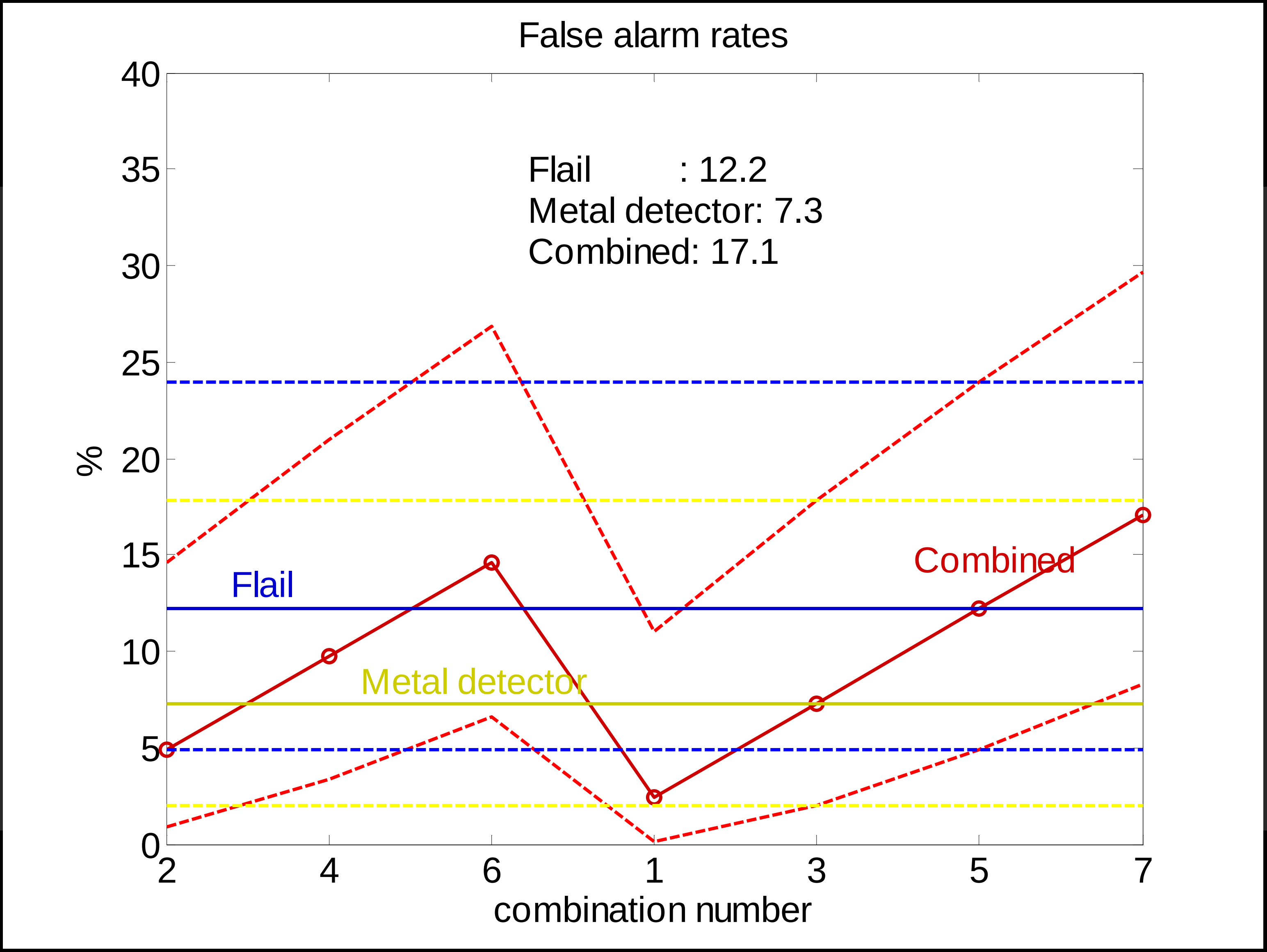
Method2 (MD): [50.4% 69.6% 84.8%]

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

Method1 (flail): $[4.9\% \quad 12.2\% \quad 24.0\%]$


Detection rates
100
90
80
70
60
50
40
30
20
10
0
%
2
4
6
1
3
5
7
combination number
Flail
Metal detector
Combined
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
2
4
6
1
3
5
7
combination number
Flail
Metal detector
Combined

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 project

They keys to a successful mine clearance system

- Use statistical learning which combines all available information in an optimal way
  - informal knowledge
  - data from design test phase
  - confounding parameters (environment, target, operational)
- Combine many different methods using statistical fusion

MineHunt System and HOSA concepts have been presented at NDRF summer conferences (98,99,01)

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

Systems: ALIS dual sensor, MD, MDD, Hydrema flail

Conclusions

danger map → clearance → update danger map

- 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 current problems

certify methods → **DeFuse** results → combine