SPATIO-TEMPORAL STATISTICS AND DATA ANALYTICS FOR EARTH SCIENCE APPLICATIONS

SUSAN C. ANYOSA

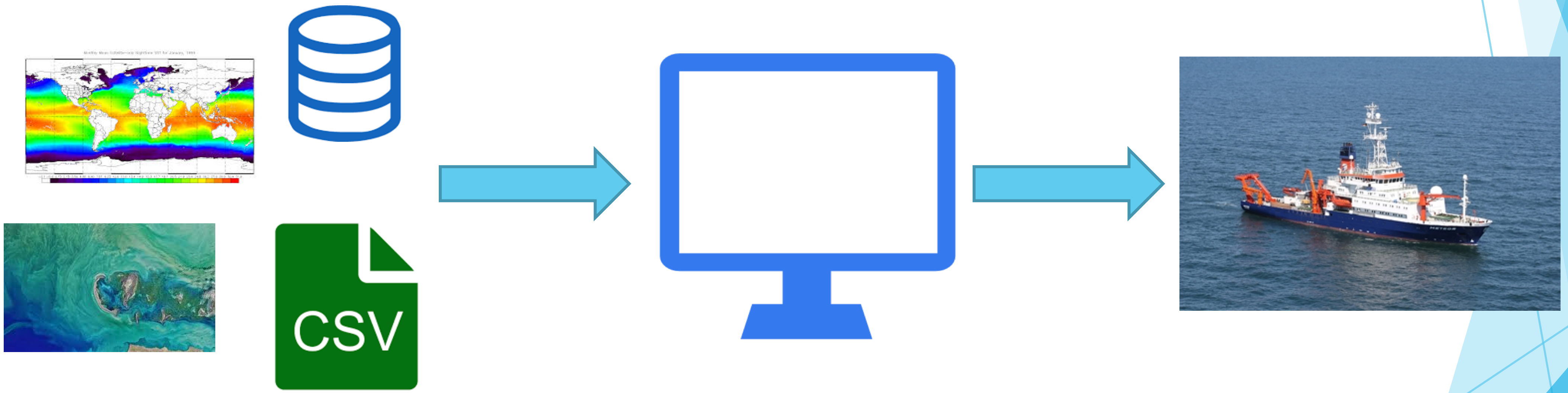
JO EIDSVIK

Objectives

- Bring together the best from modern data science, machine learning and statistical models/methods, for spatial and spatio-temporal Earth-science applications.

Motivation: The challenges to solve

How can we use our skills to solve these challenges?

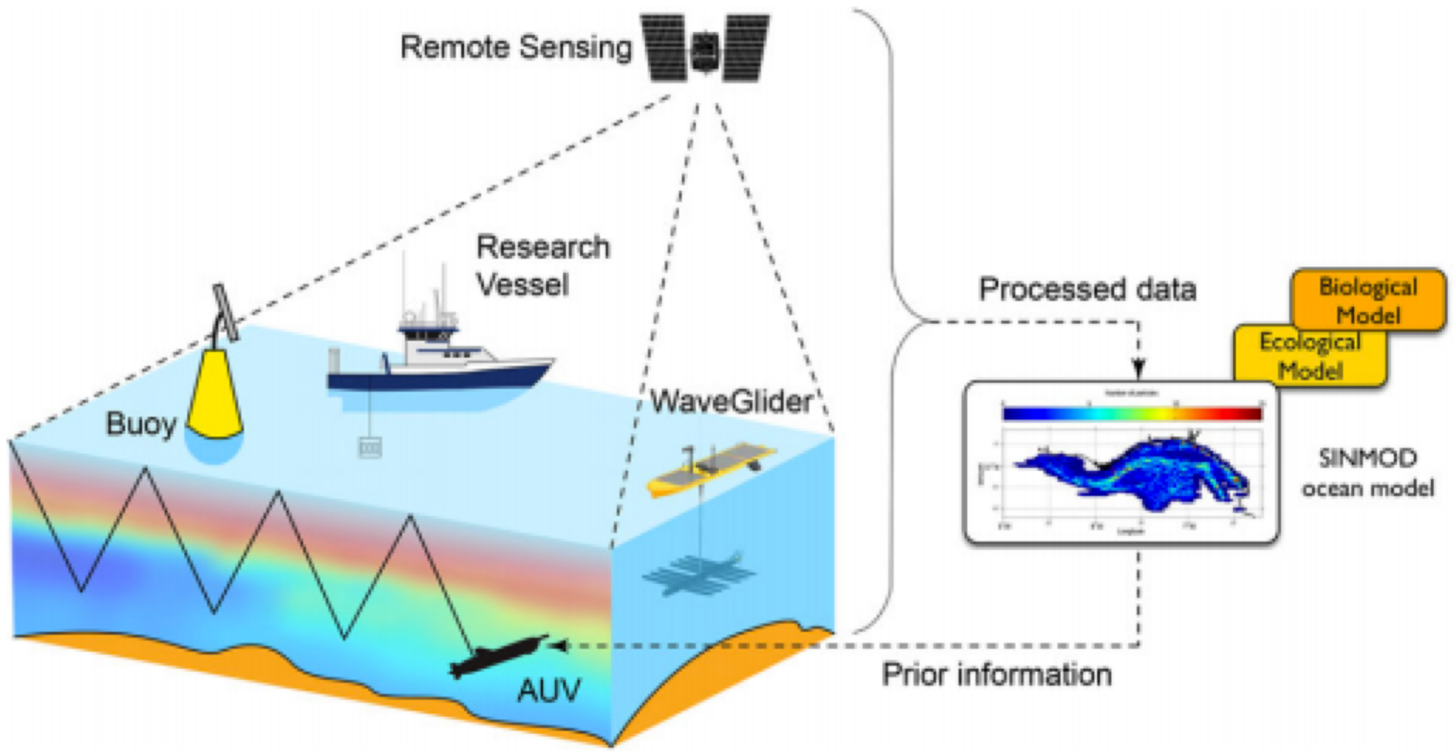


Which applications?

- Prediction of CO_2 leakage.
- Classification of sea-surface temperature (SST) images.
- Plan the design of new paths for sample expedition.

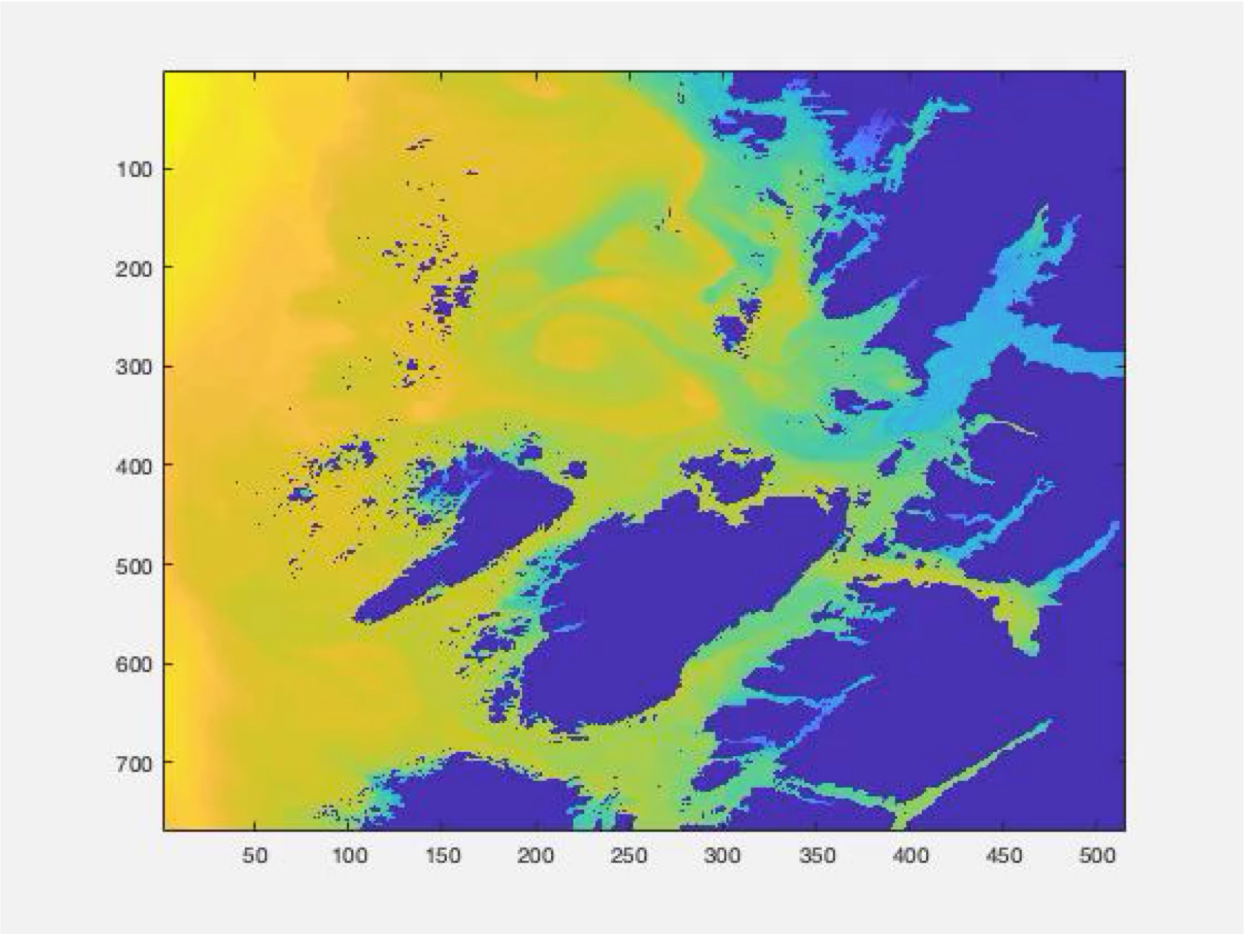
Classification of SST images

Data sources in oceanographic applications

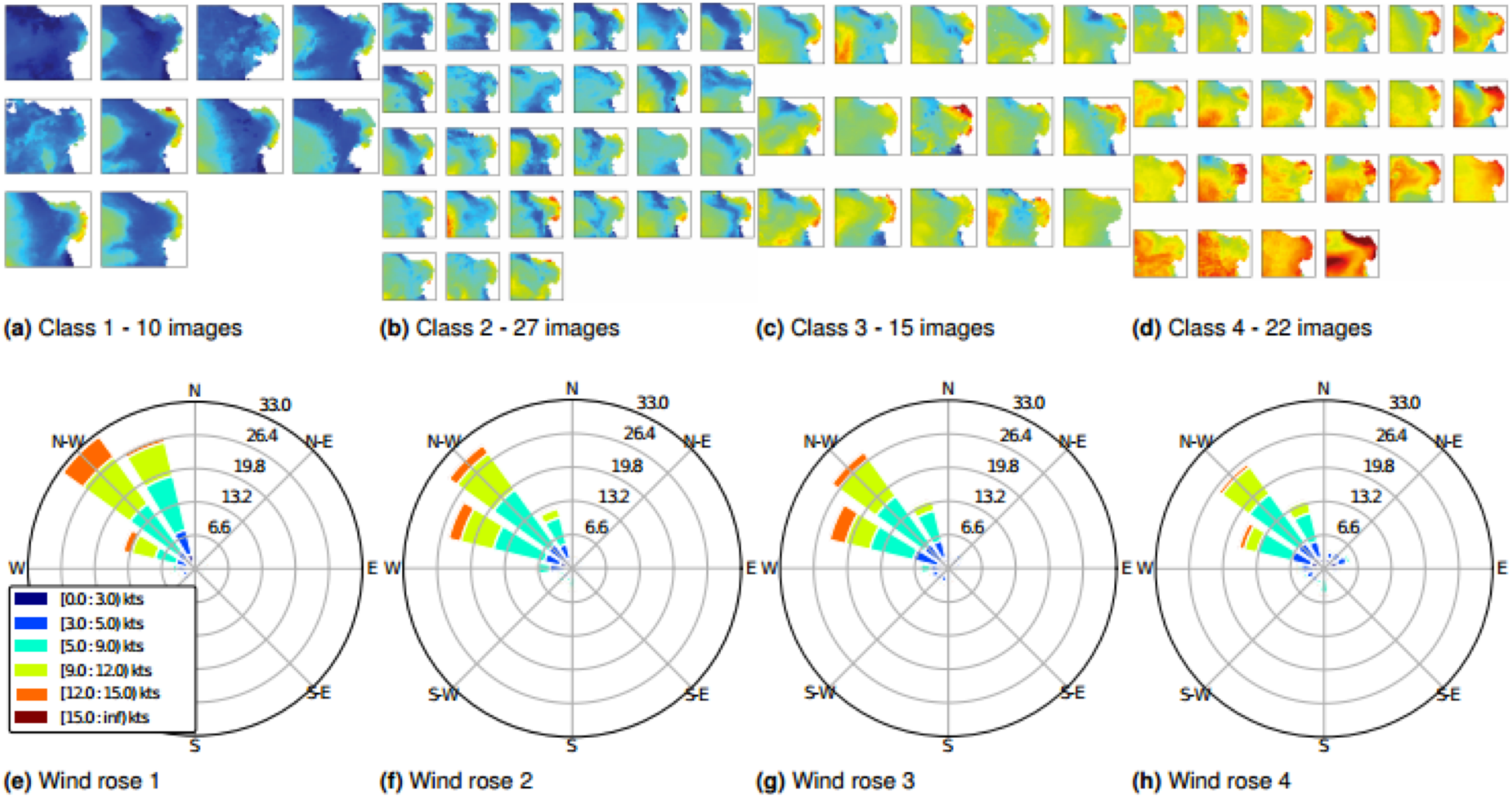


Ocean model results in space-time

- West of Trondheim fjord:



Classification of satellite data from temperature trends and variability (Monterey Bay, California)



Fossum et al. (2019), submitted to Int. Journal of Robotics research.

Approach: Gaussian processes and model-based recursive partitioning

- Fit a GP in each node.

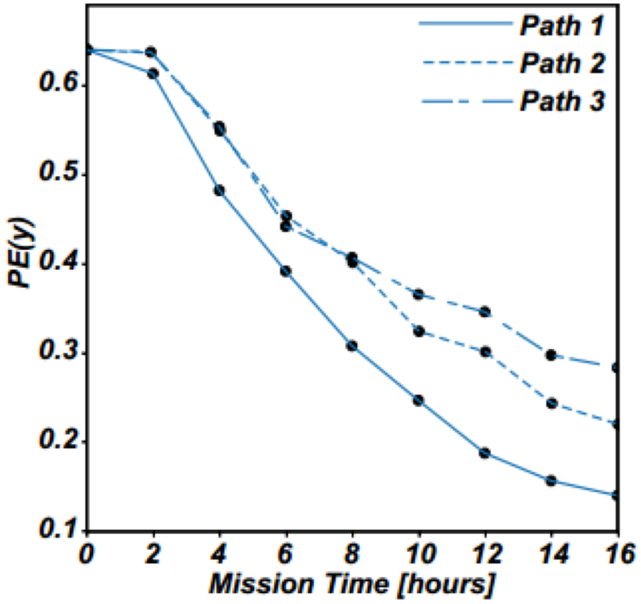
Design of new paths for sample expedition

- Prediction error

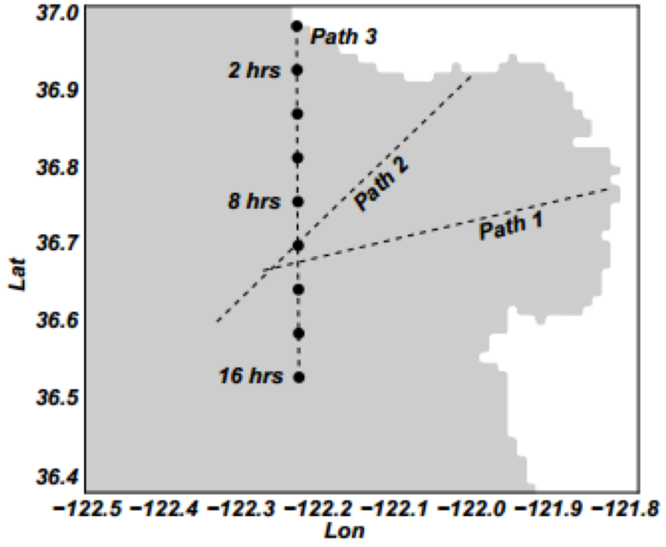
$$PE = 1 - \max_k P(x = k)$$

$$PE(y) = \int [1 - \max_k P(x = k|y)]p(y)dy$$

- What data sources are really helpful?



(a) Predicted error vs. mission time.

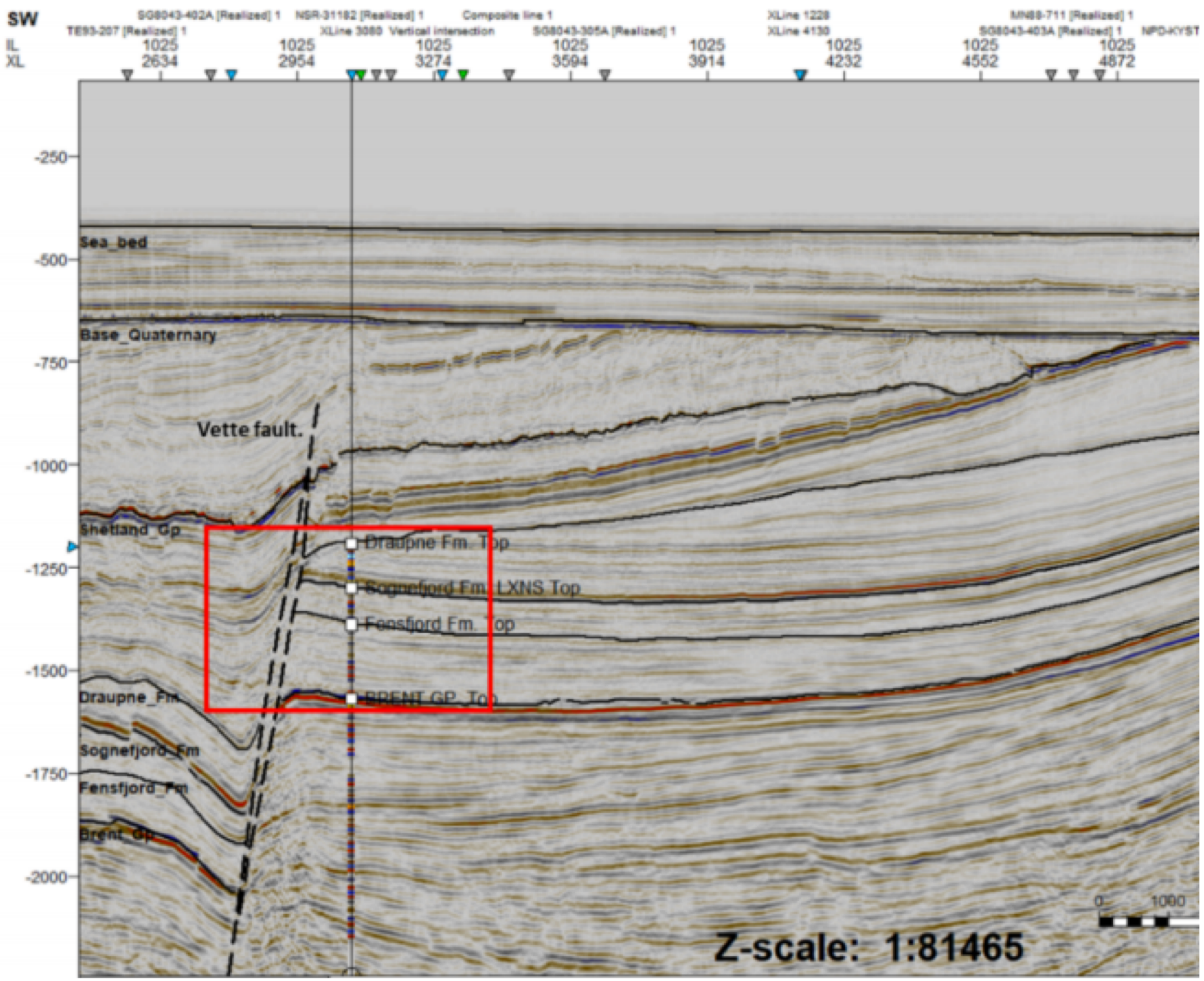


(b) The different mission paths.

Fossum et al. (2019), submitted to Int. Journal of Robotics research.

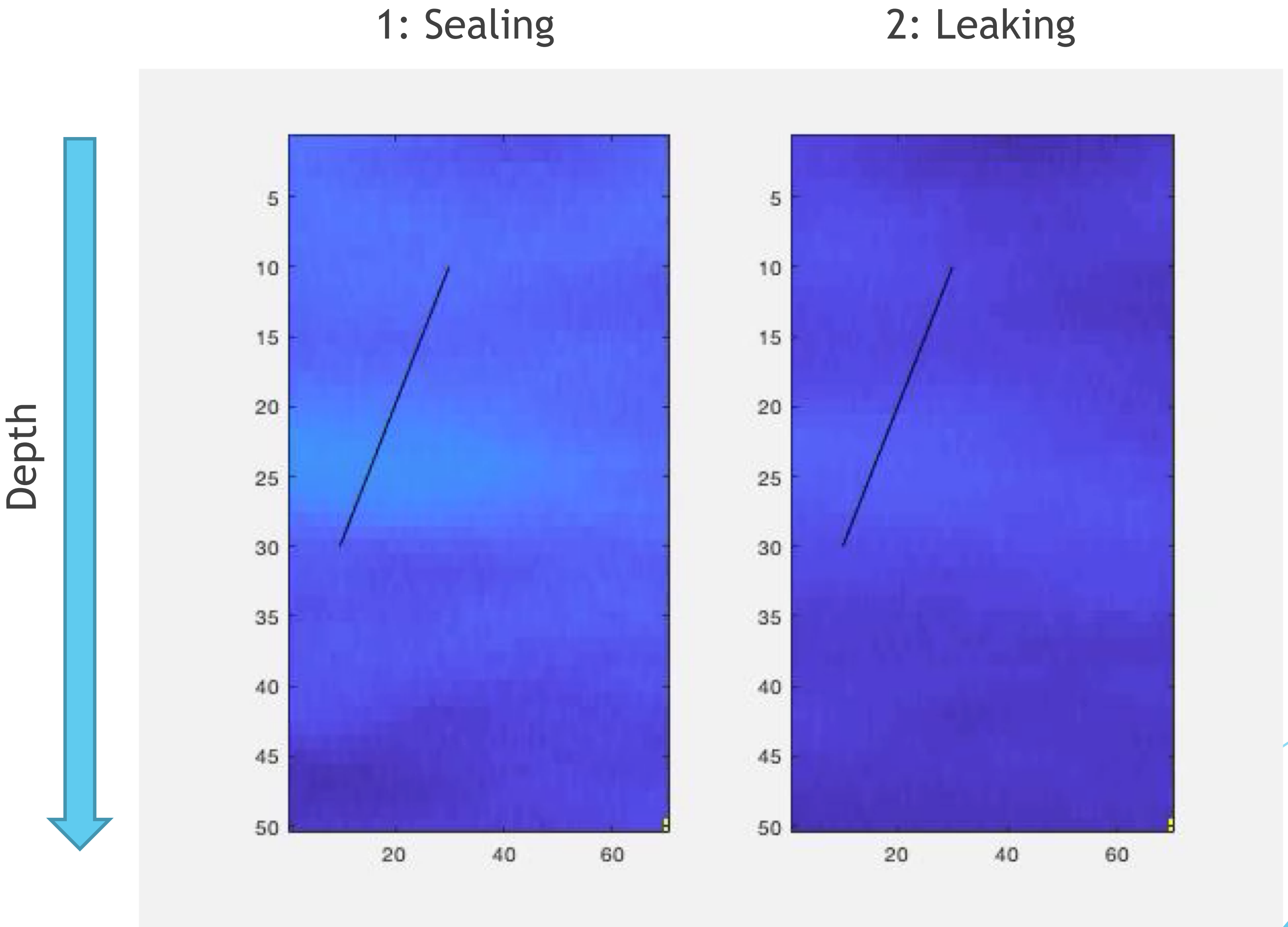
CO_2 leakage prediction

Monitoring CO_2 injection at Smeaheia



- Problem: CO_2 can leak after the injection.
- Challenge: Difficult to monitor due to seismic signatures.

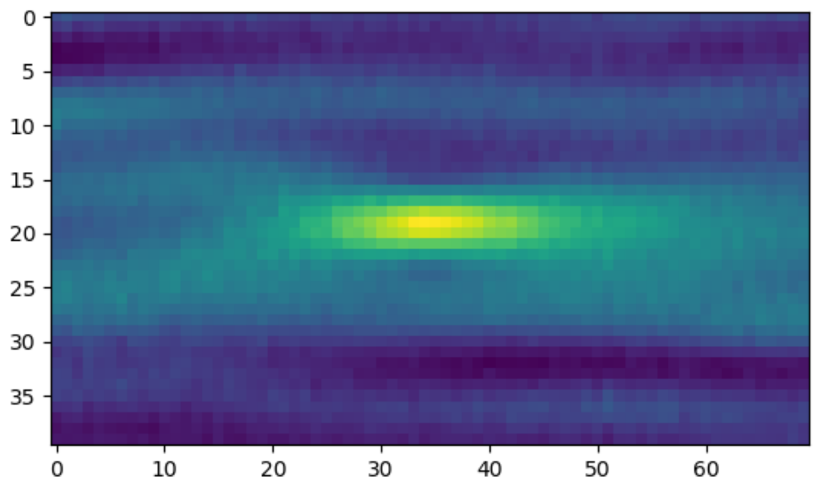
Scenarios



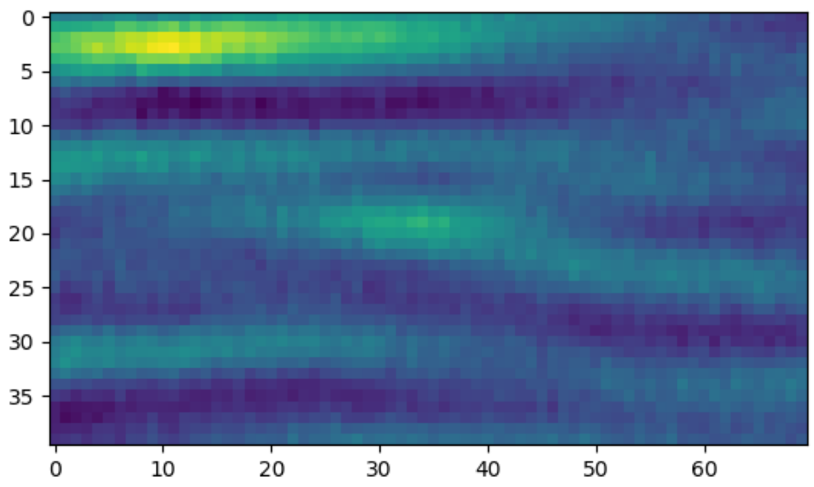
1: Sealing
2: Leaking
Depth
5
10
15
20
25
30
35
40
45
50
20
40
60
5
10
15
20
25
30
35
40
45
50
20
40
60

Approach:

- Consider 10 times of monitoring.
- For each time, simulate M images (40x70) from each scenario.
- Train in M-K images and predict the scenario in the K images of the test set.



1: Sealing



2: Leaking

- The probability of each class, conditioned on the data:

$$P(class = j \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid j)P(class = j)}{p(\boldsymbol{y})}$$

- Prior probabilities: $P(class = 1) = P(class = 2) = 0.5$

- The likelihood is approximated by a GP for each class j=1,2 with:

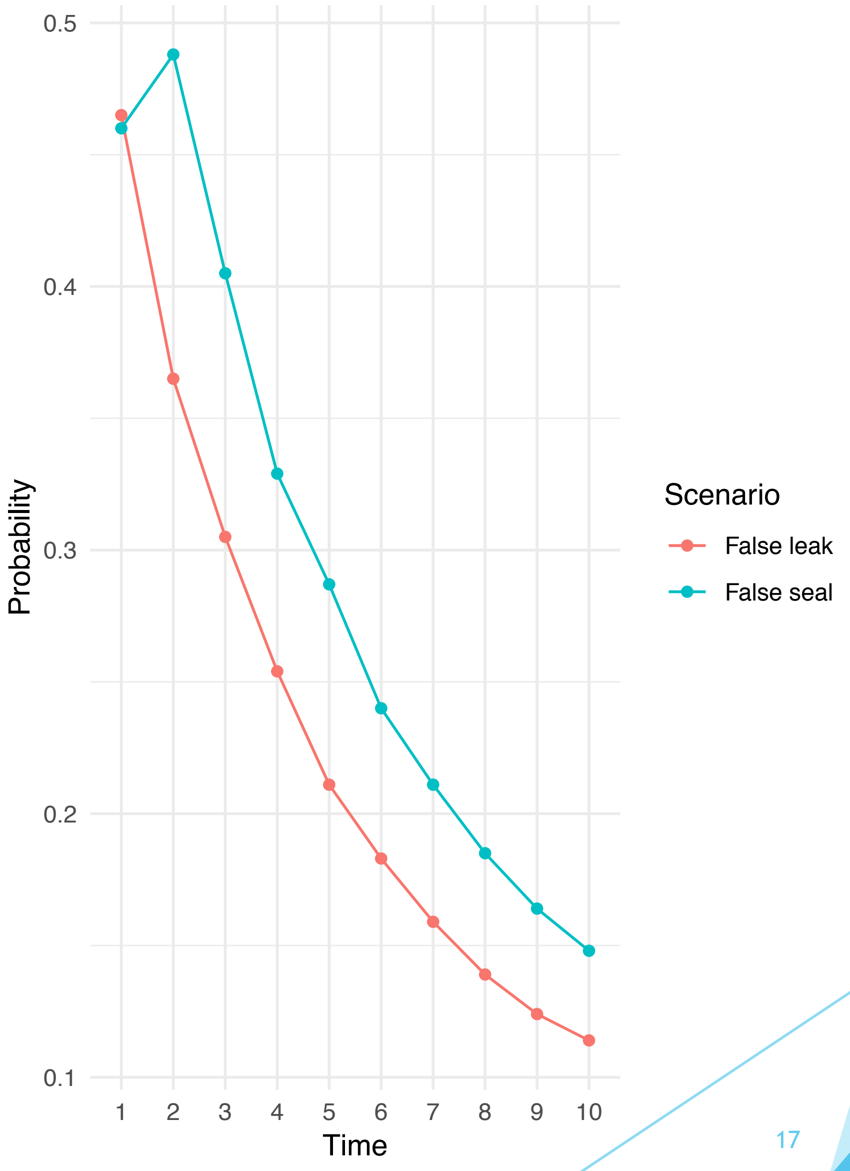
$$\hat{\boldsymbol{\mu}}^j = (\hat{\mu}_1^j, \ldots, \hat{\mu}_{2800}^j)^T$$

$$\hat{\boldsymbol{\Sigma}}^j = \begin{bmatrix} \hat{\sigma}_1^j & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{\sigma}_{2800}^j \end{bmatrix} \hat{R} \begin{bmatrix} \hat{\sigma}_1^j & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{\sigma}_{2800}^j \end{bmatrix}$$

Results

	True	
	Seal	**Leak**
Predicted Seal	True seal	False seal
Predicted Leak	False leak	True leak

- We computed a confusion matrix for each monitoring time.
- We obtained the probabilities of false leak and false seal for each monitoring time.



PhD training

- Fall 2018: MA8704 – Probability Theory and Asymptotic Techniques
- Spring 2019: MA8701 – General Statistical Methods and TDT4173 – Machine Learning and Case-Based Reasoning
- Fall 2019: MA8001 – Methods for spatial and spatio-temporal data and value of information analysis

Time Schedule