Cod Catch Forecasting through Machine Learning Algorithms at the Haul Level

Huamin Ren*, Yajie Li** and Keshav Prasad Paudel**

Huamin.ren@kristiania.no

* School of Economics, Innovation and Technology, Kristiania University College

** Faculty of Bioscience, Fisheries and Economics, UiT The Arctic University of Tromsø

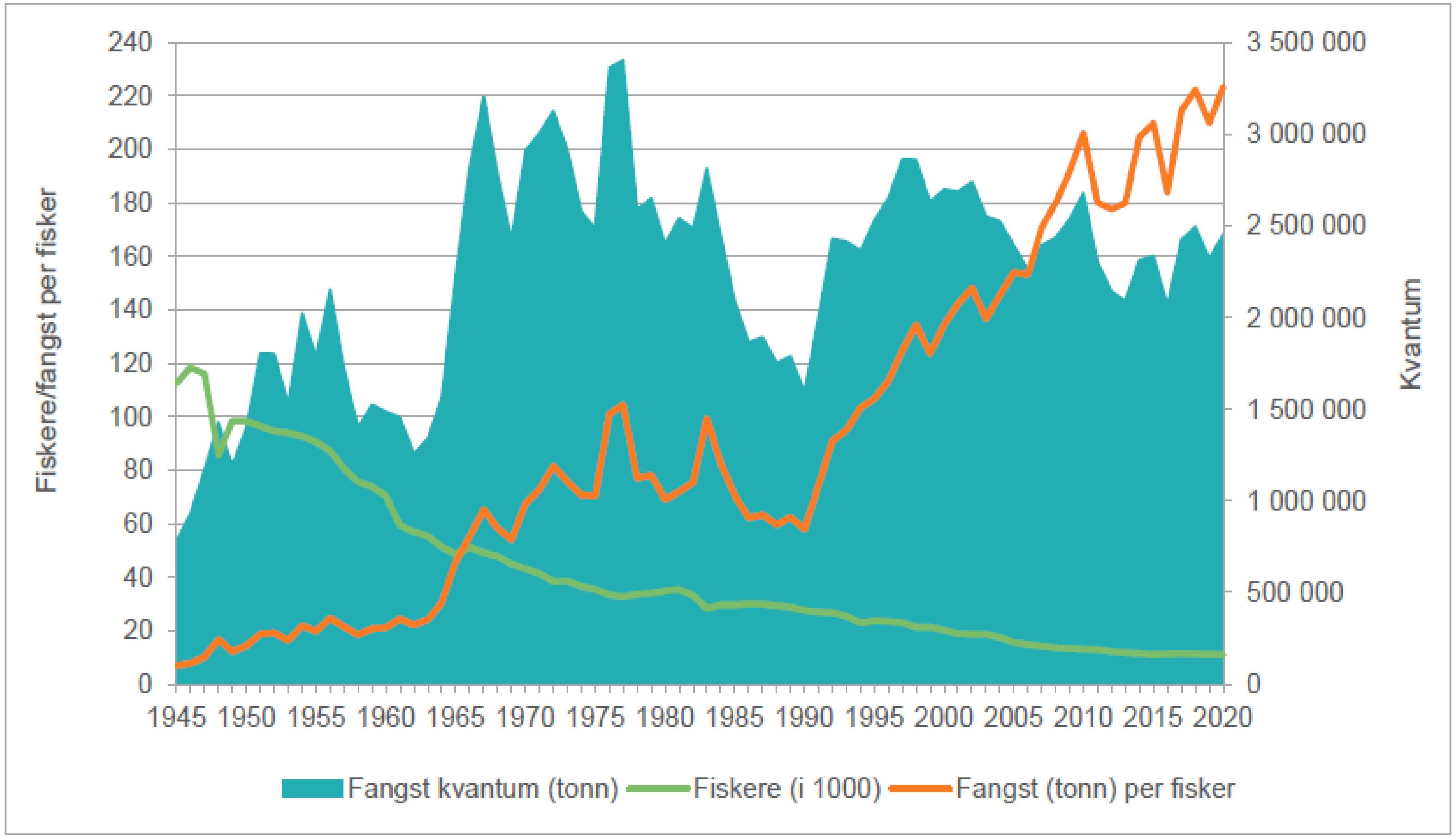
15.11.2023, Valencia, Spain

Photo: Alf Børjesson / Norwegian Seafood Council

Outline

1. Background
2. Literatures
3. Proposed method
4. Experimental results
5. Conclusion and future work

Background – Marine Fisheries in Norway



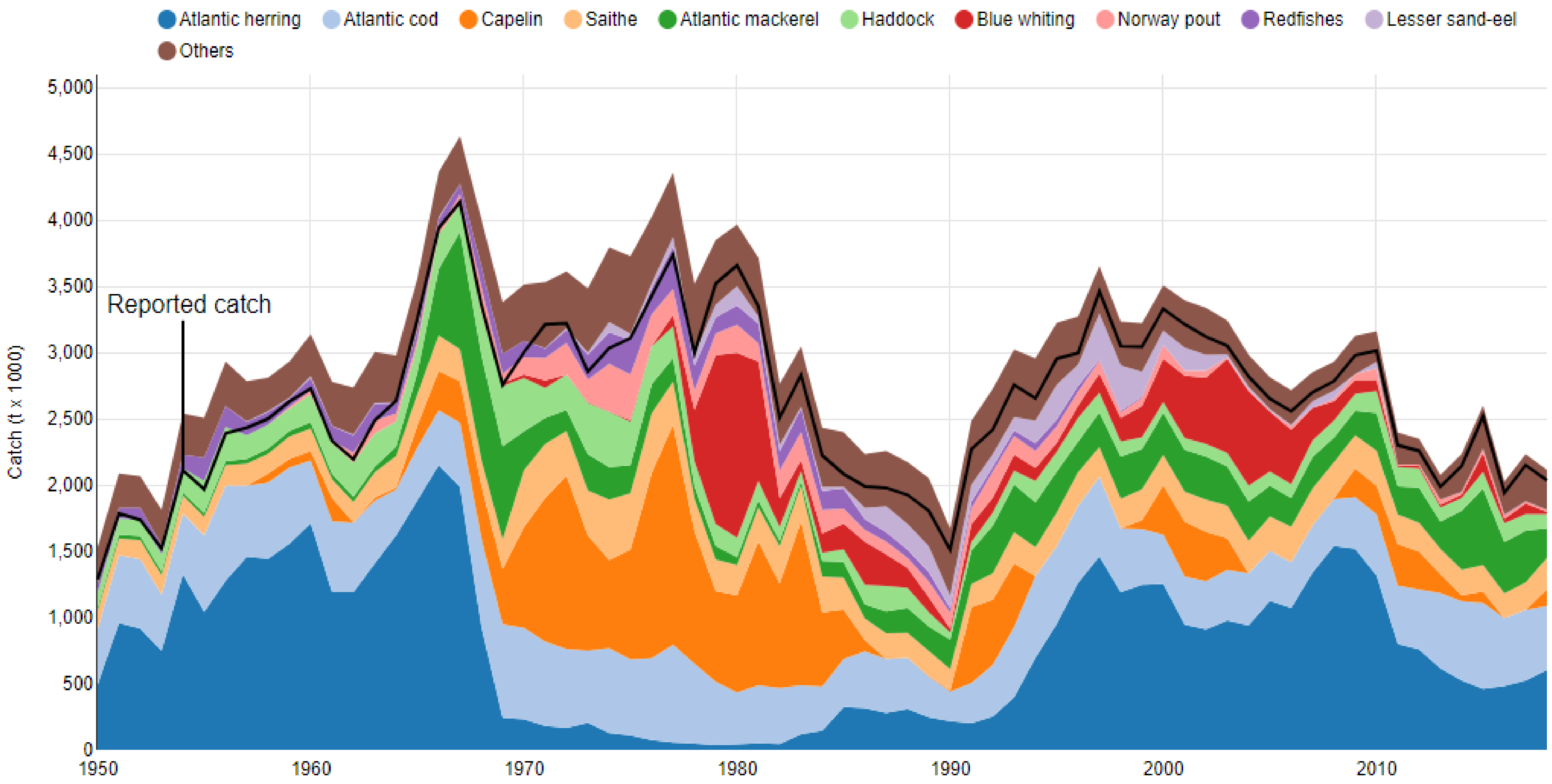
Figur 1 Fangst, fiskere og fangst per fisker 1945-2020

- 2nd contributor to the Norwegian Economy after the petroleum industry
- 2nd world's largest explorer of fish & fish products in terms of value (due to salmon)

Source: https://www.fiskeridir.no/English/Fisheries/Statistics; www.fao.org

Catches of Marine Fish Species in Norway

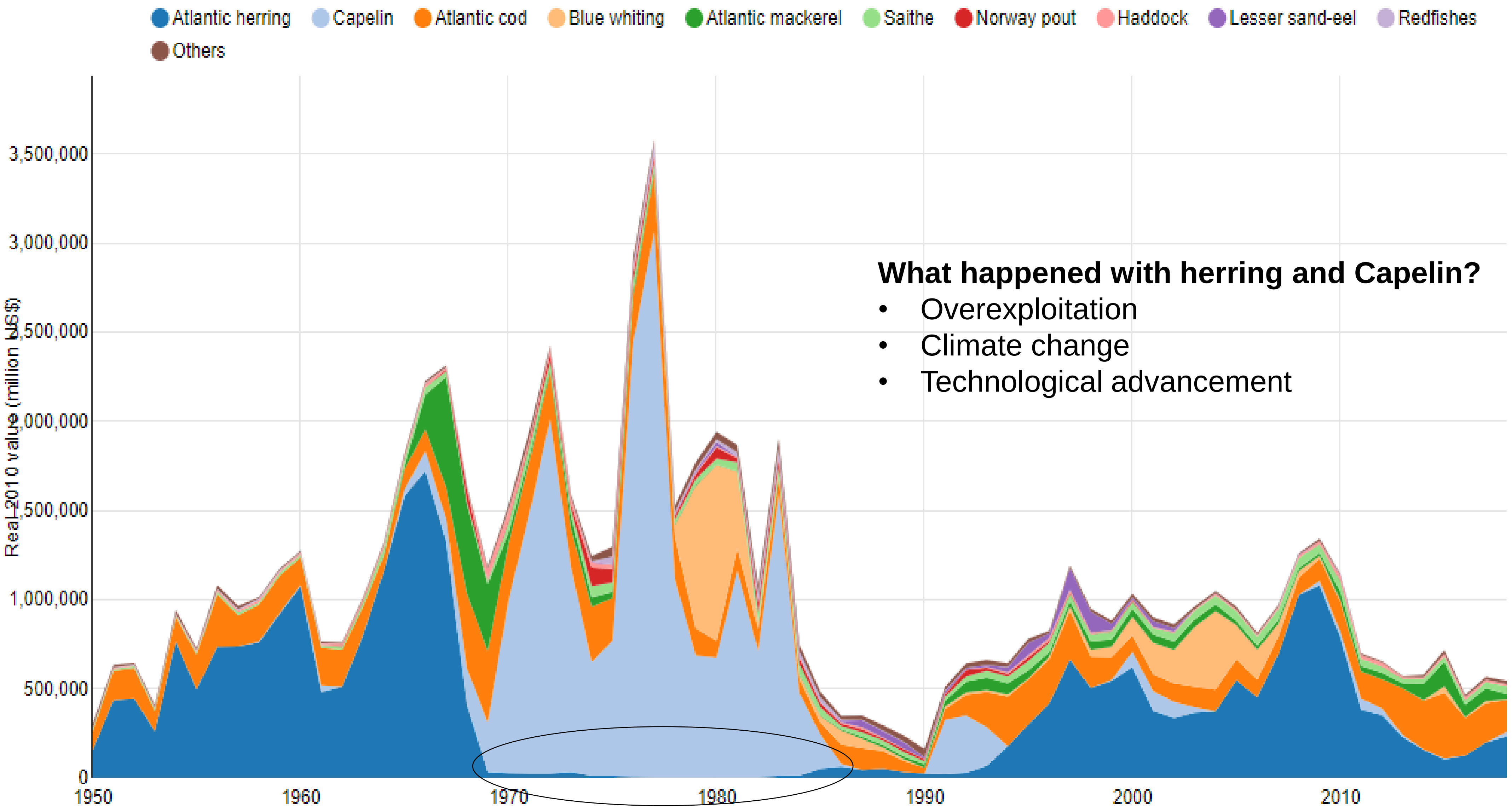
Kristiania



Source: https://www.seaaroundus.org/data/#/eez/578?chart=catch-chart&dimension=taxon&measure=tonnage&limit=10

Real 2010 Value of marine fisheries in Norway

Kristiania



What happened with herring and Capelin?

- Overexploitation
- Climate change
- Technological advancement

https://www.seaaroundus.org/data/#/eez/578?chart=catch-chart&dimension=taxon&measure=value&limit=10

Norwegian Fisheries

Based on fish species characteristics:

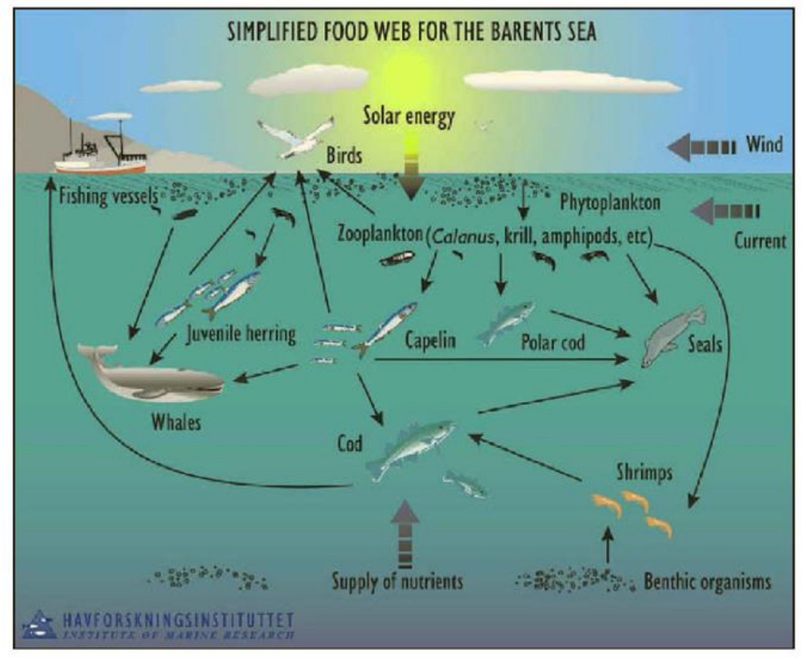
- ❖ **Pelagic species**: herring, mackerel, capelin, brisling (sprat), sandeel, Norway pout, blue whiting, etc.
- ❖ **Demersal species**: **cod,** saithe, haddock, pollack, ling, tusk, halibut

Based on fishing locations (12 miles):

- ❖ **Coastal fisheries *vs*. Ocean fisheries:**

Based on gears:

- ➢ **Conventional** gears: nets, hook-lines, seiners, pots, traps, etc.
- ➢ **Industrial** gears: trawlers, pursers, gillnets



Background

- Fishery has been an important contributor to the Norwegian Economy after the petroleum industry
- Fish catches are affected by a multitude of factors
  - fishing effort, location, types of fishing, vessels, socio-economic conditions, environmental variables etc
- Machine Learning (ML) can help fishers optimize their fishing efforts by analyzing historical catch data along with environmental factors such as ocean temperature

Literatures on Applying ML on Fisheris

- Limited…unfortunately
- Some notable attempts:
  - ➢ Predict the location of tuna fishing in the South Pacific (Zhang et al. 2022)
  - ➢ Predict marine capture fisheries and aquaculture production in Malaysia based on past production data and climate variable (Rahman et al. 2021)
  - ➢ Predict fish catches using past catches + meteorological information (Kokaki et al. 2018)
  - ➢ Assess the species richness (Leathwick, et al. 2021)
  - ➢ Analyze marine spatial planning for resolving conflicts of fisheries and other activities (Coccoli et al., 2018)
  - ➢ Estimate fishing effort allocation (Behivoke et al. 2021)
  - ➢ Evaluate fishing gear selectivity (Joshy et al., 2018)

Proposed method

- Problem Definition

Given data D:

$$D = (x_1, y_1), (x_2, y_2) ... (x_i, y_i), ... (x_M, y_M),$$

M is the number of samples in data D

x_i is the n-dimensional vector, representing the relative attributes per haul per catch

start position width, start position length, sea depth start (meters), duration - (minutes), stop position width, stop position length, sea depth stop (meters), draw distance (meters), species, round weight, etc.

Proposed method

- Objective:

$$y = f(x;\theta)$$

$$min\|\hat{y_i} - y_i\|_2, \text{ where } \hat{y_i} = f(x_i;\theta)$$

Linear Regression

RANSAC

Light GBM

Experimental results

Kristiania

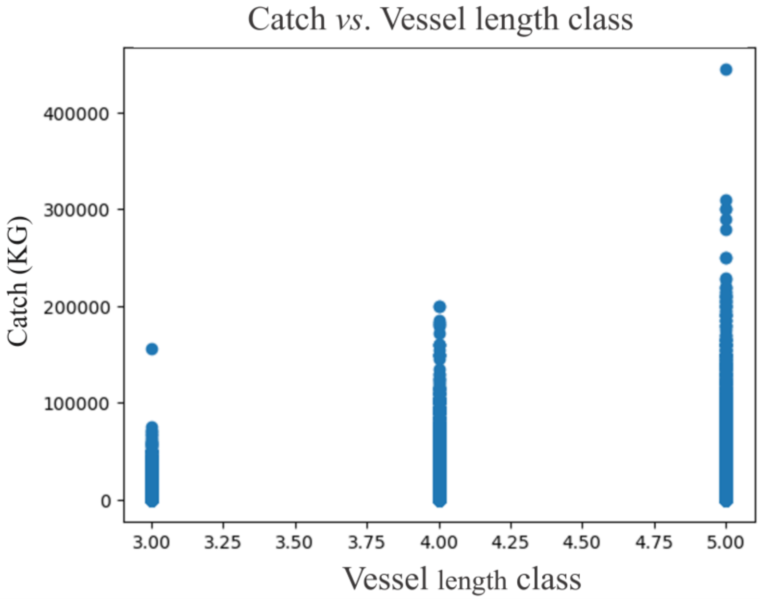
Dataset description

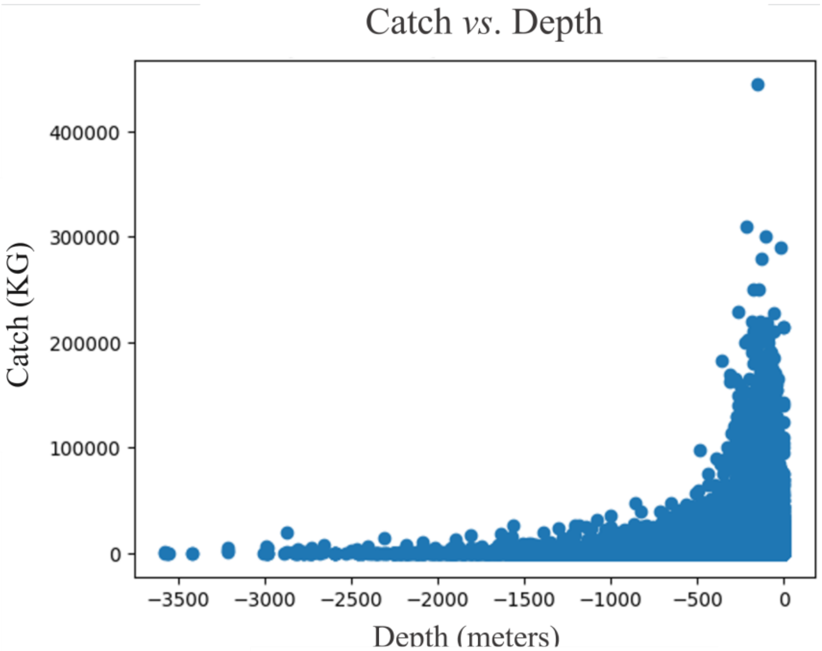
Data visualization

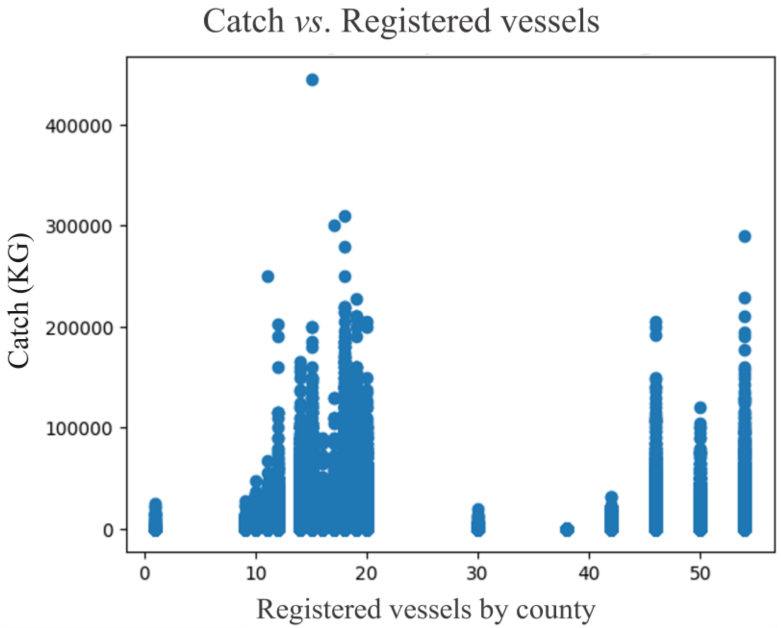
Performance and evaluation

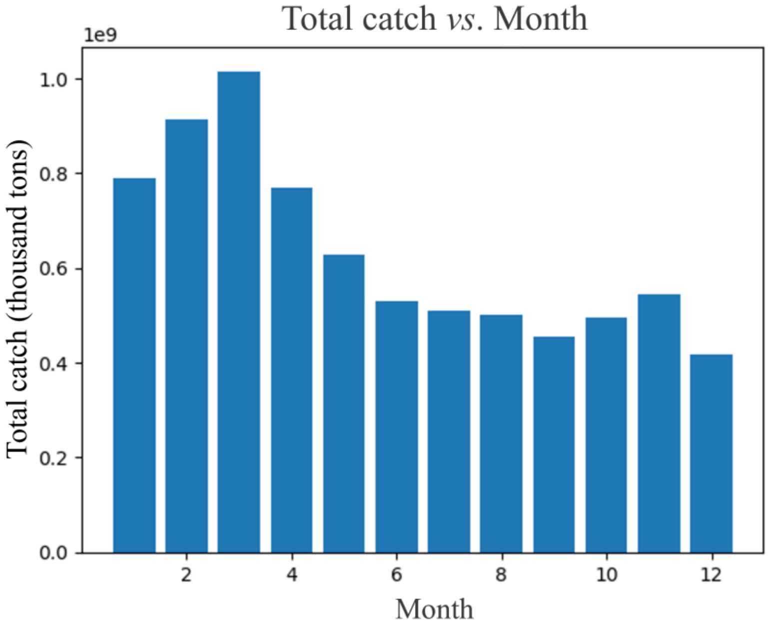
- The historical fishing data were extracted from the Vessel Monitoring System (VMS) from the Norwegian Fisheries Directorate, 2000-2022
- The dataset compromises haul time, draw distance, fishing location, catch weight, vessel characteristics, environmental variables, distance related variables etc
  - The environmental variables included two oceanographic variables: Sea Surface Temperature (SST) and sea surface Chlorophyll
  - three bathymetric and/or topographic variables: depth, slope and terrain ruggedness (rugosity)

Data visualization







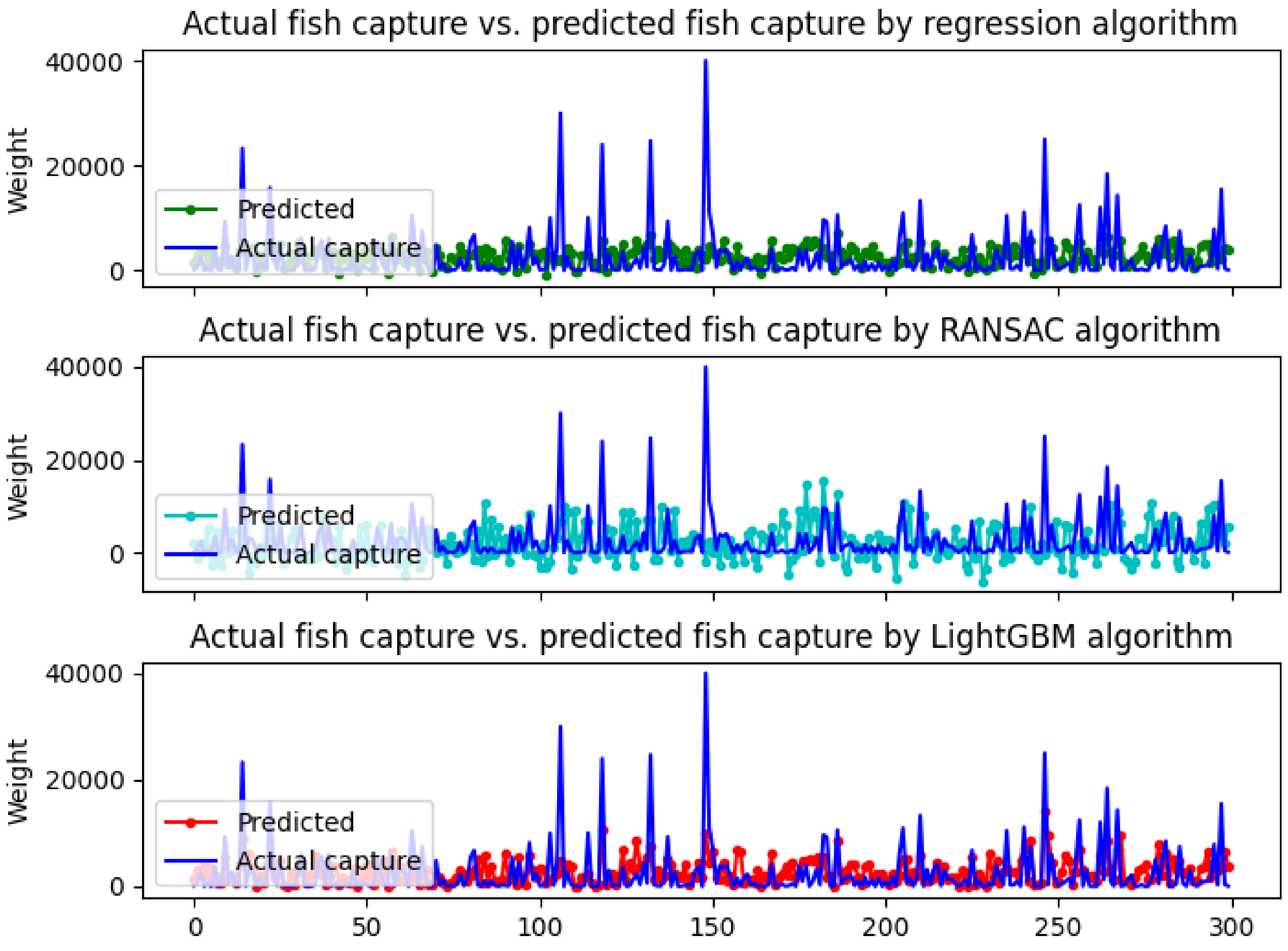


Predictions of Fish catches

Linear Regression

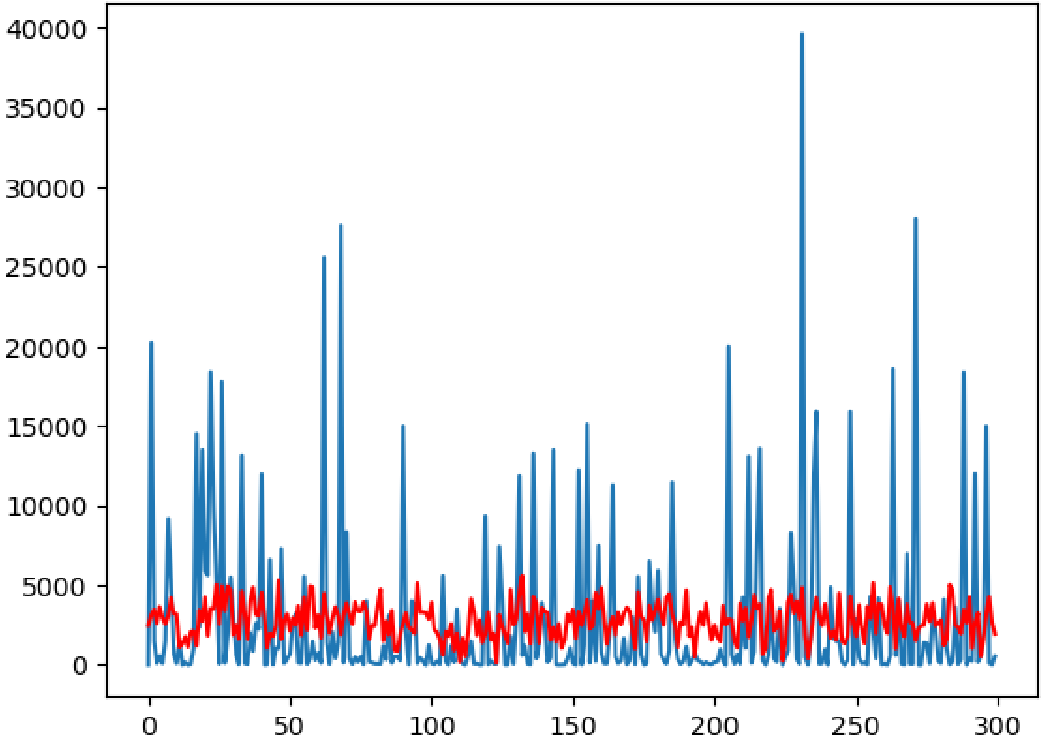
RANSAC

LightGBM algorithm

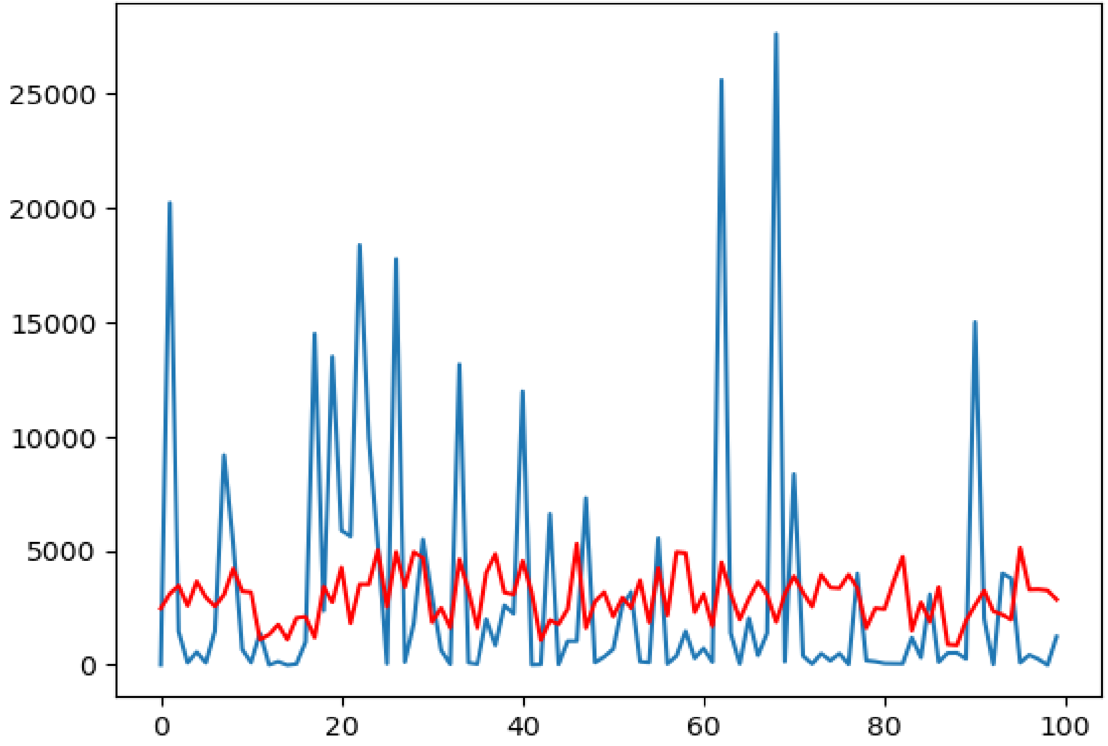


Comparing Predicted Fish Catch Performance

Linear Regression



RANSAC



Conclusion and future work

- We conducted preliminary analyses to showcase the effectiveness of linear regression, RANSAC, and LightGBM in fish catch predictions
- Model performance w.r.t large amount of noise
- Transformation of haul-level data into time series formats, targeting more vessel-focused or trajectory-driven model
- Other influencing factors, such as social-economic, policy related factors
- Expansion to other species and fisheries

- ***Huamin Ren***

- *Department of Technology, Kristiania University College
- Huamin.ren@kristiania.no

Kristiania

Thank you for your attention!