Contribution ID: **15** Type: **not specified**

Unsupervised Anomaly Detection in AIS Data from Fishing Vessels

The increasing availability of AIS (Automatic Identification System) data in the fishing sector offers new opportunities for monitoring and statistical analysis of maritime activity, but it also poses significant methodological challenges. We present an open problem motivated by the study of fishing vessels operating along the coast of Norway, aiming to promote the development and discussion of rigorous statistical approaches in a real industrial context.

The dataset integrates heterogeneous information at the vessel level: structural characteristics (length, tonnage, engine power, etc), target species, and declared catch volume. From a statistical perspective, this defines a multivariate problem with mixed-type covariates and potentially complex dependence structures. As an initial approach, we apply Isolation Forest to detect anomalous catches conditioned on vessel characteristics. This unsupervised method is scalable and flexible but raises relevant issues in industrial contexts, such as result stability, interpretability of anomaly scores, and uncertainty quantification.

Additionally, complete vessel trajectories obtained from AIS signals are available. These trajectories are treated as functional data, representing space–time curves over fishing campaigns. Functional Isolation Forest is used to identify atypical navigation patterns.

The main difficulty arises when attempting to analyze jointly both types of detected anomalies: on one hand, those related to vessel characteristics and catches and, on the other, those associated with navigation trajectories. To date, both analyses have been conducted separately. However, in an industrial context, it makes more sense to study both aspects in an integrated manner, as an anomalous catch may be linked to an unusual navigation pattern.

Therefore, it is necessary to develop a common framework that combines information from both sources. This involves integrating different anomaly measures into a single analysis framework, taking operational context into account (such as fishing area, season, or target species), and defining clear criteria for evaluating results, particularly in settings where reliable labels indicating true anomalies are unavailable.

We propose this case as an open challenge intended to foster discussion and collaboration on anomaly detection in complex and heterogeneous data. The goal is to advance toward more robust and well-founded methods that can be reliably applied in real industrial contexts, combining methodological contributions with proposals to improve result interpretation and validation.

Primary author: NAVAS-GOMEZ, Fernando (University of Granada)

Co-authors: Dr NOZAL-CAÑADAS, Rafael (Norwegian Institute of Public Health); RAYA-MIRANDA, Rocío (University of Granada)

Presenter: NAVAS-GOMEZ, Fernando (University of Granada)

Track Classification: Spring Meeting