1 Introduction

One of the most important sectors in the Norwegian economy, behind the oil sector, is the fishing industry. It creates a lot of value for Norway but it is also regulated in a way that will secure a sustainable population of fish in the oceans and also to secure a fair competition between fishing vessels. An important factor in making sure that these regulations are followed is controlling that when the vessel arrives at the port that they actually have reported the amount of fish correctly, and that the species is correct. This is done by dedicated controllers that have a limited amount of time and capacity to be everywhere and thus have to decide which port they want to control and hope that a fishing vessel arrives.

Today the controllers use their own knowledge about the reported AIS data, ERS messages and domain knowledge to estimate what port a vessel is most likely to head to. However this could be a time consuming task where making the wrong decision about which port to go to could end up costing money as the controller might be too late or be at the wrong port.

This thesis will investigate if real-time prediction of the destination port could make controls more accurate, earlier and be explainable. The investigation will be focused on finding out of deep learning on historical AIS data combined with static information about the vessel could give robust and interpretable decision support, and also importantly, that the prediction of the destination happens in a timely manner, as the controller might have to travel for a few hours.

The present thesis is also based on the work of Løvland who in a previous master thesis researched if it was possible to generate several prediction of the trajectories of a ship and infer from that which port it would arrive at, in that study it performed only slightly better than a baseline statistical model, but only if the timestep before predicting a trajectory was long enough. It also suggested for future work to include static information. Thus in this paper the goal is to investigate if this could be improved upon using a different model [1].

1.1 Problem definition

The problem can be defined as the desire to predict the destination port of fishing vessels with more accuracy than what we would get by just comparing it to statistics over previously visited ports. Since this prediction is currently done manually we would like to automate this to a larger extent.

Our research question is as follows:

> Is it possible to develop a deep learning model that can, with high precision, predict the arrival port of fishing vessels based on historical AIS data combined with vessel, catch, and route-history information?

And with sub-research questions as the following:

> 1. How much does each data source contribute to the end result? (AIS + static data, vs. AIS alone)

It is valuable to know which data sources have a high contribution to the prediction of a port. If we know that certain features barely change the result of the prediction, these could be removed from the training process, thus simplifying that process. It could also be important for Norges Råfisklag to have an even better understanding about why fishing vessels go to a specified fishing port.

> 2. What is the accuracy of the prediction based on how large a time step back in time we consider when predicting the port?

One issue when trying to predict the destination of a vessel is that we require a certain amount of data back in time to make accurate predictions about the future state. Thus we will investigate how many hours back we need to look to get an accurate prediction.

Fishing vessels are supposed to send ERS messages of type POR with information about the destination port; however, there is no legal compliance to do this and the fishing vessels face no repercussions for sending them even later. But assuming they are sent within 2 hours, this is often not enough time for the controllers to get to the port depending on where it is; thus it is important to investigate how accurate the predictions are several hours before it might arrive.

> 3. Can we also predict an accurate estimated time of arrival (ETA)?

As well as predicting the destination, it can also be useful and valuable to Norges Råfisklag and their controllers to have an estimation of when the ship will arrive. This will help them make better decisions about what time they should head to a port, to prevent unnecessary waiting at the port.

> 4. How robust is the prediction when considering potential noisy or missing data?

AIS often has a problem with the data quality as the farther away from shore you are, the higher the risk of the signals not reaching the receiver. AIS is also susceptible to spoofing attacks, which could get into our training data and cause erroneous results.

1.2 Scope

We limit the study to include only the regions of Troms and Finnmark

1.3 Limitations

1.4 Contributions

1.5 Outline

/2

Background Theory

2.1 AIS

Automatic Identifaction System (AIS) is a navigation system used by all types of vessels above the length of 15 meters, it is used for exchanging navigation signals with other ships, but also with stationary stations on ground. As the name suggests these are required to function autonomously without having to be polled by external vessels or stations. Originally AIS was designed for safety purposes of vessels, where the goal was to reduce the incidence of collisions between vessels.

Even though AIS was not designed for usage in trajectory prediction or destination prediction it still contains some useful fields that can help in those tasks. These fields are latitude, longitude, Speed over Ground (SoG) and Course over Ground (CoG), it also contains some relevant fields about the vessel characteristics that could be useful in the prediction of the destination.

What AIS actually transmits. Each AIS message contains a few types of messages [2], these are the following: *(i) Static ship data* entered at installation (name, MMSI, call sign, vessel type, dimensions); *(ii) Dynamic navigation data* which is the data that changes automatically as the vessel moces (latitude, longitude, Speed Over Ground (SOG), Course Over Ground (COG), heading, rate of turn, navigation status, time); *(iii) Voyage-related data* that is uptades as needed by the captain or similar (destination text, ETA, cargo); and *(iv) Safe-*

ty/application messages (for example addressed safety text, Aids-to-Navigation (AtoN) status). For the purpose of this thesis we will be focusing on the static ship data for the static data part and the dynamic navigation data for our time series part.

AIS uses two VHF radio channels and divides time into small slots so many ships can transmit without a central controller. Each ship's unit picks a free slot and adjusts how often it sends: more often when moving fast or turning, less often when slow or at anchor. Because VHF is line-of-sight, ship-to-ship and coastal reception typically reaches a few tens of nautical miles (around 40 nautical miles). Satellites can also receive AIS offshore, extending coverage, though in very busy areas they may miss some messages.

In Norway, AIS is integrated into a coastal monitoring network run by the Norwegian Coastal Administration (Kystverket), which receives AIS messages along the coast and creates large historical archives. National rules, aligned with EU requirements, mandate Class A AIS on larger commercial based ships and require AIS usage to fishing vessels with overall length 15 meters (including 15 meters). Foreign fishing vessels over 15 meters that land catch in Norway are also required to carry AIS. Getting access to Norwegian AIS is done by polling Kystverkets API endpoints. public use is available with restrictions (especially for vessels under 15 m and recreational crafts under 45 meters).

2.2 VMS

2.3 ERS

Fishing vessels are also required to utilize Electronic Reporting System (ERS) on their vessels. The ERS system requires the crew of the vessel to manually send messages containing information related to the potential fishing activities the vessel has done. This usage is tightly regulated and requires specific messages containing specific information and requirements for when the specific messages are to be sent. Some of these message types include the following: Departure report (DEP), the Detailed Catch and Activity report (DCA) and the Port report (POR). These messages contain mostly different information but all of the may also include information such as ID of the vessel, timestamps and so on [3]

2.3.1 POR Messages

This type of ERS message is a message that is sent as the vessel starts heading towards shore. The fishing vessels are required to send these messages two hours before landing, regardless of whether they have any catch or not. An important part of this message including the time of arrival is what port the vessel is heading to, contained in the port field of the message. They also have to report the quantity of fish they have in their cargo and also the amount of fish they plan on landing, this has to be per separate species of fish [3,4].


```json
{
  "messageType": "POR",
  "messageNumber": 784,
  "sequenceNumber": 1,

  "radioCallSign": "LNAB",
  "vesselName": "Sea Breeze",
  "registrationMark": "N-123-AB",
  "skipper": "Jane Doe",

  "messageDate": "20250923",
  "messageTime": "1015",

  "positionAtSend": {
    "latitude": "58.123",
    "longitude": "005.456"
  },

  "portOfCall": {
    "unlocode": "NOOSL"
  },

  "landingFacility": "Oslo Pier 5",
  "landingDate": "20250924",
  "landingTime": "0730",

  "etaDate": "20250923",
  "etaTime": "1830",

  "quantitiesToLand": [
    { "speciesCode": "COD", "quantity": 500, "unit": "KG" },
    { "speciesCode": "HAD", "quantity": 200, "unit": "KG" }
  ],

  "catchOnBoard": [
    { "speciesCode": "COD", "quantity": 800, "unit": "KG" },
    { "speciesCode": "HAD", "quantity": 250, "unit": "KG" }
  ],

  "cancellationCode": null
}
```


Figure 2.1: POR message

2.3.2 DEP Messages

This type of ERS message is sent out as a vessel leaves a port and contains information about the destination of fishing, the time it departs from port (estimation) and also what type of fish the vessel plans to catch. Aswell as the planned catch the vessel also have to report the amount of and type of fish they have on board at departure (if they have any catch) [3, 5].


```json
{
  "messageType": "DEP",
  "messageNumber": 123,
  "sequenceNumber": 1,
  "radioCallSign": "LNAB",
  "vesselName": "Sea Breeze",
  "registrationMark": "N-123-AB",
  "skipper": "Jane Doe",
  "messageDate": "20250923",
  "messageTime": "1405",

  "positionAtSend": {
    "latitude": "58.123",
    "longitude": "005.456"
  },

  "portOfDeparture": {
    "unlocode": "NOOSL"
  },

  "departureDate": "20250923",
  "departureTime": "1410",

  "catchOnBoard": [
    { "speciesCode": "", "quantity": 100, "unit": "KG" },
    { "speciesCode": "", "quantity": 100, "unit": "KG" }
  ],

  "plannedFishingStart": {
    "date": "20250924",
    "time": "0600",
    "latitude": "59.000",
    "longitude": "004.900"
  },

  "harvestingActivity": "TOW",
  "targetSpecies": "COD",
  "cancellationCode": null
}
```


Figure 2.2: Example departure message

2.3.3 DCA Messages

This type of ERS message contains information about the catch and activity of the vessel, if it was steaming (STE) or fishing (FIS) and more general information about when the acticity startet, GPS positions at the start and end of fishing, the quantity of fish and its species and how long the fishing activity lasted.

DCA messages are required to be sent once a day and at 23:59 UTC at the latest. If the vessel was fishing, the AC field will be set to FIS and the block B part of the message has to be sent. Some of the required content of this message is what fish was caught and the tools they used to fish it with. As well as the start and end positions during the fishing. If fishing was done several times in one day then one DCA message will contain several fishing activities.

Highlighted below is the Block A message, this type of message is to be sent by the vessel if the AC field of block A is set to STE, which means the vessel is only steaming, and performing no fishing activities [3, 6]


```json
{
  "messageType": "DCA",
  "messageNumber": 3921,
  "sequenceNumber": 2,

  "radioCallSign": "LNAB",
  "vesselName": "Sea Breeze",
  "registrationMark": "N-123-AB",
  "skipper": "Jane Doe",

  "messageDate": "20250923",
  "messageTime": "1130",

  "positionAtSend": {
    "latitude": "58.123",
    "longitude": "005.456"
  },

  "fishingPermits": [
    "PERMIT-001"
  ],

  "harvestingActivity": "TOW",

  "correctionCode": null,
  "messageVersion": 3,

  "port": {
    "unlocode": "NOOSL",
    "requiredBecauseFishingEnded": true
  }
}
```


Figure 2.3: DCA message Block A

3 Related Work

There are several studies done on the subject of utilizing transformers on AIS tracks and of course time-series data in general, usually this has mostly just been done on predicting the trajectories and not on predicting destinations directly. Even so, there might be relevant knowledge to be had based on that as well.

3.1 Predicting the destination port of fishing vessels utilizing transformers

The study by Løvland researches how to predict the destination port of Norwegian fishing vessels by using a transformer to generate trajectory forecasts several hours in the future, and then converting tracks that end close to a port into a probability based on how many tracks hit certain ports. The region of interest they looked at is Troms and Finnmark, with AIS data between 2016 and 2023 that are combined with ERS events in the same time range to separate out steaming segments after fishing stops and before arrival at a port. The tracks are cleaned by dropping points with implausible speeds and a long time between messages, interpolated to 5 minute steps, and normalized for latitude, longitude, SOG and COG. For the forecasting they utilized a type of transformer called the TrAISformer which utilizes a four hot encoding and a categorical cross entropy loss that is suited to multi modal routes, the predicted paths are smoothed before port inference. The authors utulize a growing proximity radius (2 to 5km) plus a low speed ner port logic that can differentiate true arrivals from vessels that just pass by, and the end result is a top k port probabilities from multiple sampled futures. The modesl is compared to a simple statistic based baseline model that utilizes just the historical destinations of the ports to give a probability of the most likely port.

Empirically, the study found that longer input time steps improves both the trajectory error and also the downstream port prediction accuracy. Reported metrics include top 1 and top 3 port accuracy, with invalid cases where no predicted path enters any port radius. The trained model was also not to time consuming to train, with a training time of around two hours to train on the stated hardware, and around nine seconds to produce 16 futures for a single input. Chosen hyperparameters (five minute sampling, 16 trajectories, 100 prediction) correspond to about an eight hour look ahead. The authors also highlight some limitations such as dependence on recent AIS, region specificity. But propose some ideas for future work such as adding static vessel and route features and reframing the task as a direct classification of ports instead of inferring from a trajectory.

The paper by Løvland etal is relevant to this master thesis since it is looking at the same task, with the same dataset and region of interest. Thus the findings in this paper offers some concrete preprocessing choices and a suggestions for future work that we will be investigating in this paper. [1]

3.2 AIS Data-Driven Maritime Monitoring Based on Transformer: A Comprehensive Review

The paper "AIS Data-Driven Maritime Monitoring Based on Transformer: A Comprehensive Review" is a survey paper on how transformer architectrues are used for the purpose of AIS based maritime monitoring, organizing prior work into three subjects: vessel trajectory prediction, vessel behavior detection and vessel behavior prediction. The paper argues that transformers are well suited because the self attention mechanism helps in capturing long range dependencies in AIS sequences and handles long inputs more efficiently than recurrent models. The review also summarizes recurring AIS data issues, outliers, missing values, and reception gaps and describes some preprocessing practices seen across the literature, including filtering erroneous positions/speeds/courses and gap filling strategies that some studies adopt before modelling.

Beyond comparing methods, the authors also compiled and cleaned a public AIS dataset drawn from the studies they reviewed, totalling around 640 million messages from 19016 vessels across six ship categories. They provide descriptive statistics by vessel type, noting differences between voyage lengths across classes, and present the dataset as a resource to support future transformer based research on AIS. In the task overviews, the survery distinguishes between generative trajectory forecasting and classification style forecasting (predicting discretized future states or regions), and it outlines behavior oriented applications ranging from anomaly detection to state prediction of vessel activity. Finally, the paper also mentions some future dtirections, ehmphasizing the need for improved handling of data quality and trajectory reconstruction, as well as multi source data fusion (incorporating environmental information) and it recommends exploring different transformer variants and hybrids to better capture interactions in complex traffic. Overall the paper contributes to a good overview of different ways of using transformers for the purpose AIS research. And it provides a curated dataset and some guidance on data preprocessing [7]

3.3 Multi-path long-term vessel trajectories forecasting with probabilistic feature fusion for problem shifting

The paper adresses long horizon (up to 12 hours) vessel trajectory forecasting utilizing only one to three hours of AIS history by reframing prediction as a trajectory reconstruction over likely routes to reduce uncertainty. The study area is the Gulf of St. lawrence, where ships follow multiple plausible route; to capture route level behavior the ocean is discretized into equal hexagonal cells. The method augments standard AIS spatiotemporal inputs with probabilistic features, a classifier proposes likely route polygons and a likely destination cell, which are then combined into the learning pipeline. Cyclic transforms of longitude, latitude, and bearing are used to stabilize the learning process on angular/periodic variables. The model uses parallel dilated convolution blocks for spatial cues, followed by a bidirectional LSTM with position aware attention and a decoder to produce future coordinates. The data used is AIS data between 2015 and 2020 focused on cargo and tanker vessels.

The authors attained a F1 score of about 0.95 for route prediction and a 0.75 F1 for destination on held out data, end to end forecasting reports mean 11km and median 6km positional error with an $R^2 > 98\%$. Compared against transformer baselines, the proposes models are smaller and faster to train while also matching or exceeding accuracy on turns and route choices especially. Ablation testing indicate the main gains come from the probabilistic features, cyclic encondings and attention emphasizing recent timesteps. The trained model is deployed within the smartWhales decision system to help reduce vessel whale encounters in atlantic Canada. [8]

3.4 Vessel Destination Prediction Using a Graph-Based Machine Learning Model

This study investigated the vessel destination prediction problem as a multiclass link prediction on a heterogeneous maritime graph built from AIS voyages. Ports are retrieved from the world port index, arrivals/departures are inferred from stop behavior and speed threshold to reconstruct the voyages. The graph contains vessel nodes and port nodes; with edges representing voyages and also carrying information such as month of departure, draught, cargo type, and the weight at departure time. The model learns node and link embeddings via a random walk with a word2vec style training, then deriving link embeddings for vesel, port pairs that is then fed to a k nearest neighbors (KNN) classifier for infering the destination. The authors experiments used AIS data from Danish waters between the January 2014 to March 2021 only looking at tanker ships. The graph model was implemented using StellarGraph and scikit-learn.

The authors write that this graph based approach outperforms many classical baselines, such as logistic regression. KNN, random forests and CatBoost on accuracy, precision, recall and F1, and adding graph topology features further improve the precision and F1. A comparison against manually endtered AIS destination fields show that many entries are incorrect, while the model does attain a higher correctness, most residual errors cluster on the low frequency ports, with a stronger performance on often visited destinations. The paper also notes some limitations, such as a regional scope, possible class imbalances, possible overfitting to vessel identifiers and points to future work on global data and graph neural networks [9].

This information provides some useful knowledge for the present study as it also is directly focused on destination prediction rather than trajectory regression, thus it could motivate the testing of a graph based model as well.

3.5 Vessel Trajectory Prediction Based on AIS Data: Dual-Path Spatial–Temporal Attention Network with Multi-Attribute Information

The paper studies the task of trajectory prediction for maritime vessels using AIS data. The authors extract the motion patterns and then predict future tracks; they also perform bench marking against several deep kearning and graph bases baselines. Their own model DualSTMA runs two attention paths (temportal to spatial, and spatial to temporal) and explicitly fuses multi-attribute information, where they seperate dynamic inputs from static inputs using a gating mechanism. For the inputs they use latitude, longitutede, speed, heading, and various static fields such as vessel type and size. Their study is designed for pure trajectory prediction but not a final destination prediction, with evaluation in degrees/meters over horizons up to 50 minutes.

On U.S. coastal AIS data (2021) the method the authors propose outperforms LSTMs, GRUs, Transformers STGAT and METO-S2S across Average Displacement Error (ADE), Final Displacement Error (FDE), and Mean Absolute Error (MAE). The reported gains are large, especially at longer horizons (approximately a 69% improvement in 50 minute ADE/FDE vs the best baseline), and qualitative ploits show stronger performance on turns and interaction heavy scenes. The authors also run ablations showing that modeling both temporal dynamics and spatial interactions matter and that treating static attribnutes with a gate is better than concatenation.

Compared to the goals of this thesis, the paper tackles a closely related problem, that is predicting where the track will go rather than predicting which port it will reach. Still, the AIS features overlap with the usage of the dynamic and static information, thus this paper gives a strong rationale for the usage of both dynamic and static information in an attention based sequence model [10].

3.6 Temporal Fusion Transformer

The paper introduces the Temporal Fusion Transformer (TFT), an attention based architecture for multi horizon forecasting that combines a strong preidction accuracy with built in interpretability. TFT is designed to handle both static covariates, known future inputs and historical observed inputs within one model. The core elements of the model include a gating mechanism (to give less importance to irrelevant components), instance wise variable selection networks for static/past/future inputs, static covariate encoders that produce context vectors of the static data, local tmeporal processing via a sequence to sequence layer and an interpretable multi head self attention decoder for long range dependencies. the attention model shares value projections across heads and averages attention weights to help aid interpretability. while outputs are quantile forecasts trained with a quantile loss.

The experiments ran by the authors span from electricity, traffic, retail and financial volatilty datasets. TFT reports good results on P50 and P90 risk across horizons, and ablations highlight major contributions from the local temporal processing and self attention, with further gains from static encoders, variable selection and gating. The authors also demonstrate its interpretability uses cases, global variable importance, persistent temportal patterns, and regime shift detection, and note practical training details such as employing a single interpretable attention layer and feasible single GPU training times. [11]

3.7 On the Exploration of Temporal Fusion Transformers for Anomaly Detection with Multivariate Aviation Time-Series Data

- **Goal:** Explore a forecasting-based anomaly detection approach for aviation using the *Temporal Fusion Transformer* (TFT), trained to learn nominal behavior and flag anomalies as large forecast errors; proof-of-concept focuses on *unstable approaches (UA)*.

- **Data:** Multivariate time-series from MITRE's threaded track + digital flight data, plus weather/context (runway ID, head-/crosswind, wind dir/speed, visibility, wind–runway diff).

- **Sampling windows:** Last 20 minutes before touchdown (240 steps at 5 s); each sample uses a 64-step look-back (320 s) and 8-step look-ahead (40 s), yielding 169 overlapping samples per flight.

- **TFT bits that matter:** Variable Selection Networks (VSNs) provide global feature-importance scores; gating (GRNs) and static covariates enhance interpretability and performance.

- **Feature sets tested:** *TFT-1* (tracks + runway), *TFT-2* (+ headwind/crosswind), *TFT-3* (+ wind dir/speed, visibility, wind–runway diff). *TFT-2* generally gives the lowest speed-forecast error (mean RMSE 3.06).

- **Baseline comparison:** A univariate "ARIMA-like" TFT (speed→speed) performs worst (mean RMSE 3.43), indicating multivariate inputs add value.

- **Single vs multi-output:** Predicting speed and altitude jointly is feasible, but single-output models are more accurate (e.g., speed RMSE 3.06 single vs. 3.13–3.27 multi; altitude mean RMSE 83.48 single vs. 88.72–96.77 multi).

- **Feature selection via VSNs:** Training with top-ranked features (*TFT-select*) matches *TFT-2* overall (mean RMSE 3.02 vs 3.06) and can reduce compute without hurting accuracy.

- **Anomaly signal:** Use RMSE time profiles; Fisher's discriminant identifies strongest nominal–UA separation 26 timesteps before touchdown, guiding where to threshold.

- **Detection performance:** At that point, speed-only TP=38.63

- **Takeaway:** Label-efficient, interpretable forecasting of nominal behavior can surface UA precursors and localize timing; broader datasets, tuning, and richer outputs could further improve detection. [12]

Bibliography

[1] A. B. Løvland, H. Fredriksen, and J. M. Bjørndalen, "Predicting the destination port of fishing vessels utilizing transformers," *Maritime Transport Research*, vol. 8, p. 100131, Mar. 2025. Article ID 100131.

[2] "AIS Dynamic data and Static Information." Imarinex [Online]. Hentet fra: https://www.imarinex.com/ais-dynamic-data-and-static-information/ (Lastet ned: 27. okt. 2025).

[3] "Elektronisk rapportering (ERS)." Fiskeridirektoratet [Online]. Hentet fra: https://www.fiskeridir.no/yrkesfiske/rapportering-paa-havet/elektronisk-rapportering-ers (Lastet ned: 27. okt. 2025).

[4] "Forskrift om posisjonsrapportering og elektronisk rapportering for norske fiske- og fangstfartøy (ERS-forskriften)." [Online], 2009. Hentet fra: https://lovdata.no/dokument/SF/forskrift/2009-12-21-1743/KAPITTEL_3#%C2%A713.

[5] "Forskrift om posisjonsrapportering og elektronisk rapportering for norske fiske- og fangstfartøy (ERS-forskriften)." [Online], 2009. Hentet fra: https://lovdata.no/dokument/SF/forskrift/2009-12-21-1743/KAPITTEL_3#%C2%A711.

[6] "Forskrift om posisjonsrapportering og elektronisk rapportering for norske fiske- og fangstfartøy (ERS-forskriften)." [Online], 2009. Hentet fra: https://lovdata.no/dokument/SF/forskrift/2009-12-21-1743/KAPITTEL_3#%C2%A712.

[7] Z. Xie, E. Tu, X. Fu, G. Yuan, and Y. Han, "Ais data-driven maritime monitoring based on transformer: A comprehensive review," *arXiv*, 2025.

[8] G. Spadon, J. Kumar, D. Eden, J. van Berkel, T. Foster, A. Soares, R. Fablet, S. Matwin, and R. Pelot, "Multi-path long-term vessel trajectories forecasting with probabilistic feature fusion for problem shifting," *Ocean Engineering*, vol. 312, p. 119138, Nov. 2024. Article ID 119138.

[9] R. Gouareb, F. Can, S. Ferdowsi, and D. Teodoro, "Vessel destination prediction using a graph-based machine learning model," in *Network Science (NetSci-X 2022)* (P. Ribeiro, F. Silva, J. F. Mendes, and R. Laureano, eds.), vol. 13197 of *Lecture Notes in Computer Science*, pp. 80–93, Springer, Cham, 2022.

[10] F. Huang, Z. Liu, X. Li, F. Mou, P. Li, and Z. Fan, "Vessel trajectory prediction based on ais data: Dual-path spatial–temporal attention network with multi-attribute information," *Journal of Marine Science and Engineering*, vol. 12, p. 2031, Nov. 2024. Art. no. 2031.

[11] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.

[12] B. Ayhan, E. P. Vargo, and H. Tang, "On the exploration of temporal fusion transformers for anomaly detection with multivariate aviation time-series data," *Aerospace*, vol. 11, p. 646, Aug. 2024. Art. no. 646.