Annals of Operations Research (2025) 345:533–568
https://doi.org/10.1007/s10479-024-06457-9

ORIGINAL RESEARCH

Arctic sea ice thickness prediction using machine learning: a long short-term memory model

Tarek Zaatar[1] · Ali Cheaitou[1] · Olivier Faury[2] · Patrick Rigot-Muller[3]

Received: 30 January 2024 / Accepted: 20 December 2024 / Published online: 8 January 2025

Abstract

This paper introduces and details the development of a Long Short-Term Memory (LSTM) model designed to predict Arctic ice thickness, serving as a decision-making tool for maritime navigation. By forecasting ice conditions accurately, the model aims to support safer and more efficient shipping through Arctic waters. The primary objective is to equip shipping companies and decision-makers with a reliable method for estimating ice thickness in the Arctic. This will enable them to assess the level of risk due to ice and make informed decisions regarding vessel navigation, icebreaker assistance, and optimal sailing speeds. We utilized historical ice thickness data from the Copernicus database, covering the period from 1991 to 2019. This dataset was collected and preprocessed to train and validate the LSTM predictive model for accurate ice thickness forecasting. The developed LSTM model demonstrated a high level of accuracy in predicting future ice thickness. Experiments indicated that using daily datasets, the model could forecast daily ice thickness up to 30 days ahead. With monthly datasets, it successfully predicted ice thickness up to six months in advance, with the monthly data generally yielding better performance. In practical terms, this predictive model offers a valuable tool for shipping companies exploring Arctic routes, which can reduce the distance between Asia and Europe by 40%. By providing accurate ice thickness forecasts, the model assists in compliance with the International Maritime Organization's Polar Code and the Polar Operational Limit Assessment Risk Indexing System. This enhances navigation safety and efficiency in Arctic waters, allowing ships to determine the necessity of icebreaker assistance and optimal speeds, ultimately leading to significant cost savings and risk mitigation in the shipping industry.

Keywords Neural networks · Long short-term memory (LSTM) · Climate forecasting · Regional forecasting · Ice thickness · Northern sea route (NSR)

✉ Olivier Faury
ofaury@em-normandie.fr

1 University of Sharjah, Sharjah, United Arab Emirates

2 EM Normandie, Métis Lab, Le Havre, France

3 Maynooth University, Maynooth, Ireland

1 Introduction

Due to considerable ice melts, navigating along the Russian Arctic shores, particularly the Northern Sea Route (NSR) along the northern coasts of Russia between Murmansk and Bering Sea (Koyama et al., 2021), has gained in attractivity with possible economic interests (Fu et al., 2018; Yang et al., 2024) accompanied by greenhouse gas (GHG) emissions and time savings (Lindstad et al., 2016). Developing the different zones of a country or a region is usually a priority for the authorities, enabling resource exploitation (Di et al., 2023). Operating shipping routes in the Arctic Ocean can, in principle, reduce the navigational distances between Europe and Asia compared to those of the Suez Canal Route by around 40%, as shown in Fig. 1, offering considerable fuel savings and thereby reducing CO_2 emissions (Lasserre, 2014; Cheaitou et al., 2023). Moreover, the Arctic is a region with vast potential opportunities for commercial activities for the oil and gas industry, tourism, and mining, and growing commercial activity means that the Arctic is a vital market arena for both shipowners and their industry partners (Borch et al., 2012).

Yet the presence of ice makes navigation dangerous for vessels and expensive while posing risks to the environment (Zhang et al., 2020). Indeed, one of the possible risks of using the

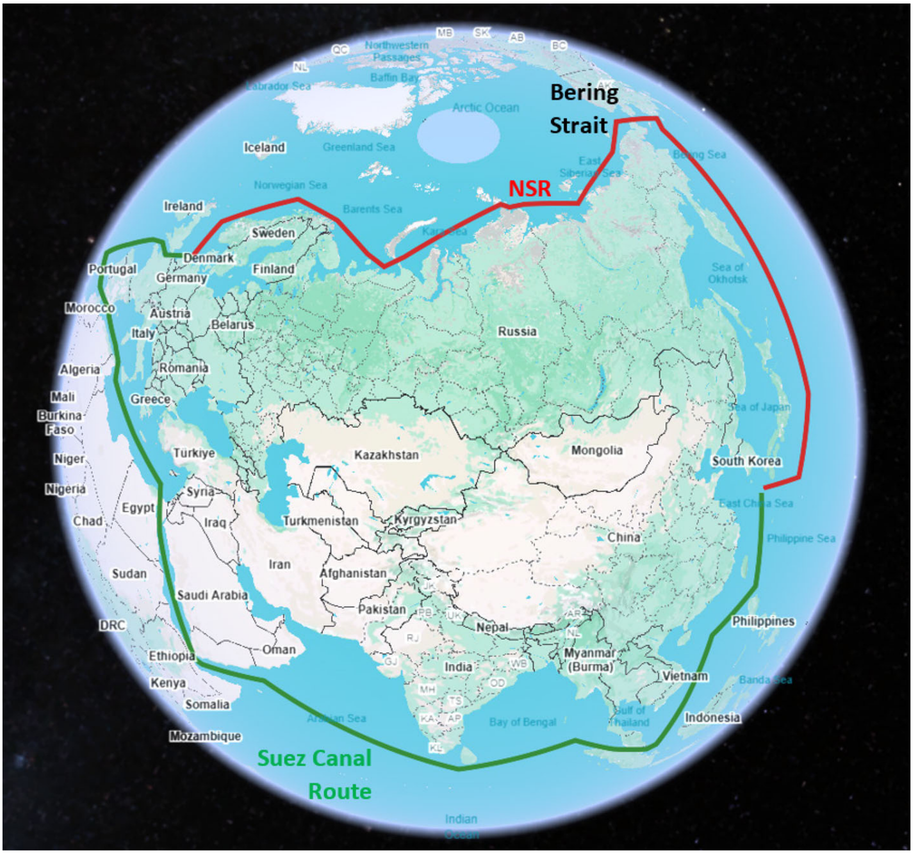


Fig. 1 The Northern Sea Route compared to the Suez Canal Route (Based on Google Earth)

NSR lies in the uncertainty of the length of the navigation season due to the ice conditions in the Arctic (Fu et al., 2018). Hence, ice-thickness in the Arctic is a seasonal parameter that varies between summer and winter due to changes in solar irradiance and temperature.

As highlighted by the Polar Code (IMO, 2014), ice is one of the main threats to navigation impacting decisions taken by desk officers, underwriters, or shipping companies (Fedi et al., 2018a, 2018b; Fu et al., 2016; Zhang et al., 2020) at both the strategic and operational levels (Zeng et al., 2014; Rigot-Muller et al., 2022) with economic and ecological consequences (Cariou et al., 2019; Cheaitou et al., 2022; Li et al., 2024). Thus, being able to anticipate the ice conditions that the vessel may encounter, i.e., ice extent and ice thickness, is paramount for stakeholders sailing through the Arctic (Fu et al., 2016; Zhang et al., 2020).

As stressed by the National Snow Ice Data,[1] the extent of the ice is difficult to predict from one year to another even if the general trend is negative (Meier & Stroeve, 2022). However, unlike sea ice concentration, for which real-time data can be recorded, it is very hard to have accurate real-time data on sea ice thickness in the Arctic since this requires having tools physically implemented in the sea and comprehensive satellite readings. Instead, the available (historical) data is used to simulate and assimilate predictions of ice thickness to achieve values that are as close as possible to what the real values would be. Indeed, CryoSat2 satellite ice thickness records mainly originate from the inadequate knowledge of snow depth and density of snow/ice (Liu et al., 2019). Other researchers have examined predicting Arctic Sea ice thickness, concentration, and temperature using various techniques such as data assimilation, software simulations, and linear and multi-linear regressions (Ali et al., 2023; Diebold et al., 2021; Li et al., 2024).

Hence predicting ice conditions is necessary for the shipping companies that are willing to provide a service in the Arctic in order to make decisions related to the investments in winterized vessels, the assessment of the sailing costs, the estimation of the transit time. All such elements depend on ice conditions (Cariou et al., 2019; Fu et al., 2016; Lasserre, 2014).

Yet, if the extent of ice is an essential parameter (Li et al., 2024; Zhang et al., 2020), the thickness also has to be considered as it is the case with Polar Operational Limit Assessment Risk Indexing System (POLARIS). POLARIS is a decision-making tool integrated within the Polar Code aiming to define the capacity of a vessel to sail in defined ice conditions based on the level of winterization of the vessel (IMO, 2016). Additionally, to provide navigation as safe as possible, underwriters require shipowners to define the route likely to be used (Fedi et al., 2018a, 2018b).

The relationship between ice thickness and speed has already been demonstrated (Faury et al., 2020; Fu et al., 2016; Zhang et al., 2020) and is foremost in planning a voyage within icy waters from risk management and economic points of view. First, from a risk management perspective, the presence of ice may also impact the rescue response time and act as a catalyst in case of an incident or accident. Another major risk that results from harsh weather and ice conditions is shipping accidents that may involve oil spills and contamination of the Arctic environment (Aksenov et al., 2017; Zhang et al., 2020). Finally, numerous accidents are due to inadequate speed (Marchenko, 2014; Zhang et al., 2020).

Second, the speed of the vessel directly impacts fuel consumption and thus the cost of the journey (Cariou et al., 2019). Besides, ice conditions make the sailing time to reach the destination longer, and therefore the costs are higher. Furthermore, the combination of both transit time and cost greatly influences the attractiveness of Arctic shipping lines (Cheaitou et al., 2020) with slower sailing speed, vessels will need longer travel times, which adds more cost and reduces the number of trips performed per year (Theocharis et al., 2019). Moreover,

[1] https://nsidc.org/sea-ice-today/sea-ice-tools/charctic-interactive-sea-ice-graph

the Arctic is a sensitive area where GHG and non-GHG emissions have a larger impact than in other areas (Faury et al., 2020). Indeed, as temperatures are rising and the global warming phenomenon in the world is becoming more tangible, the Arctic Sea's ice has been changing in recent decades in terms of thickness and concentration, especially during the summer period; this has made the Arctic Sea of high interest in different environmental, business, and logistics aspects, and in particular due to its rich environment with many oil and gas fields and other natural resources, especially in the western part of the Russian Arctic (Fu et al., 2018).

The literature has examined the changes in the Arctic Ocean ice conditions extensively, and many articles have raised the need for an accurate prediction of the Arctic Sea's ice parameters, namely ice thickness. According to Kozmenko et al. (2018), the average temperatures of the Arctic have almost doubled over the past 100 years, resulting in a rapid decline of the multiyear ice and the reduction of the ice thickness and quantity. These conditions contribute to improving the likelihood of having better navigation along the NSR, making it one of the feasible shipping routes that provide tremendous shipping benefits (Lee & Song, 2014). However, rising temperatures and disappearing ice during the summer season increase the risk of ships encountering drifting ice, which constitutes another reason to need to predict ice thicknesses. Moreover, it is also estimated that 22% of the world's undiscovered oil resources are in the Arctic, 84% of which is projected to be offshore (Milaković et al., 2014). These changes have motivated the development of the Arctic transportation system (Khrapov & Yushchenko, 2019), and recent research studies have leaned toward having the NSR as the preferential route for use in the Arctic (Theocharis et al., 2018).

On the other hand, an extensive review of the capability of Artificial Intelligence (AI) tools and in particular Artificial Neural Networks (ANN) to produce precise climate and weather predictions has been proven (Aichouri et al., 2015; Li et al., 2024). Moreover, the applicability of ANN in predicting nonlinear weather data has been examined by Abhishek et al. (2012). More specifically, LSTM models and their extensions have been used for different climate-related predictions such as wind (Wang et al., 2024a, 2024b), temperature and humidity (Yang et al., 2023), solar irradiance (Qing & Niu, 2018), weather forecasting (Karevan and Suykens, 2020; Venkatachalam et al., 2023), snowmelt flood (Zhou et al., 2023), and visibility (Ortega et al., 2023). Moreover, LSTM has been used in other climatic-related contexts such as the prediction of carbon emission-related economics (Shahzad et al., 2023) and carbon market (Sadefo Kamdem et al., 2023). Other authors have also developed machine learning models for different prediction objectives related to weather such as rainfall (Umamaheswari & Ramaswamy, 2024), heatwaves (Bhoopathia et al., 2024), or atmospheric parameters such as air quality or ozone concentration (Madan et al., 2025; Wang et al., 2024a, 2024b). On the other hand, ANN-based approaches have been used to investigate regional economic dynamics in low-carbon contexts (Di et al., 2024a, 2024b). More generally, the impact of the digital economy and digital innovations on sustainability, regional development, and energy have been investigated by different studies (Sun et al., 2024; Wang et al., 2025, Xue et al., 2024).

AI tools have been used for the analysis and prediction of sea ice thickness in different studies. For instance, ANN ensemble (the outputs of a set of separately trained ANNs combined to form one unified prediction) has been used in the prediction of ice thickness on several Canadian lakes during the early winter ice growth period (Zaier et al., 2010). The results of an ANN ensemble have been promising and showed an improvement in terms of result accuracy (Zaier et al., 2010). Another study by Wang et al. (2017) used a Convolution Neural Network to analyze Synthetic Aperture Radar (SAR) images in estimating sea ice concentration in the Gulf of St. Lawrence on the east coast of Canada. Additionally, the added value of Convolutional Neural Network has been highlighted by Li et al (2024). Ressel et al. (2015) investigated the possibility of using an ANN to classify the ice type from data provided by SAR images. LSTM has been used to predict the break-up date of river ice in Heilongjiang province in China (Liu et al., 2023) and for the ice of St. Marys River and the Laurentian Great Lakes in North America (Liu et al., 2022). In addition, Adaptive Weighted Ensemble Learning has been used to predict ice coating (Guo et al., 2024). Moreover, LSTM has been used specifically to develop a ship-following model, designed to ice-covered waters, and used to predict ship-following behavior (Duan et al., 2024).

More specifically, many studies have focused on Arctic Sea ice parameters (Diebold et al., 2021) and different tools have been used to estimate and predict them. In particular, LSTM models have been used to forecast Arctic snow depth (Dong et al., 2022), monthly Arctic Sea ice extent (Ali et al., 2023), total monthly pan-Arctic Sea ice extent (Wei et al., 2022), and Arctic Sea ice concentration (Li et al., 2021; Liu et al., 2021; Phutthaphaiboon et al., 2023; Zheng et al., 2022). More specifically, Zaatar et al. (2021) have developed an LSTM model to predict sea ice thickness and applied it on a very limited area in the Arctic. Wu et al. (2024) used factor selection and machine learning methods to correct the bias in the available Arctic ice thickness data products.

It is worth mentioning that LSTM models are particularly effective for predicting Arctic Sea ice thickness since they are designed to capture complex temporal dependencies in sequential data (Li et al., 2024). Unlike simpler models such as AutoRegressive Integrated Moving Average (ARIMA) or Prophet, which rely on linear assumptions and are best suited for stationary data or seasonal trends, LSTMs excel at modeling nonlinear relationships and retaining information over long time lags, making them suitable for the intricate nature of sea ice dynamics (Box et al., 2015). In comparison to traditional ANNs, which do not inherently account for temporal dependencies, LSTMs have an advantage due to their specialized memory cell structure, which allows them to retain vital information over extended sequences (Hochreiter & Schmidhuber, 1997). This is crucial for modeling the long-term and seasonal interactions between factors including temperature, wind, and ocean currents that influence sea ice changes LSTMs hence offer a more robust approach for accurately predicting Arctic Sea ice thickness compared to ARIMA, Prophet, or traditional ANNs. Table 1 provides a comparison between LSTM and the main time series models (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

It is worth noting that most of the works in the literature have focused on predicting the sea ice concentration, and few have been directed toward the prediction of sea ice thickness. Moreover, to the best of the authors' knowledge, none of the studies have used LSTM models to predict NSR sea ice thickness and its surroundings, which this paper will try to do. Equally to the best of the authors' knowledge, the existing literature has not covered large Arctic expanses in terms of ice thickness prediction and with a high level of accuracy, which constitutes an essential research gap given the importance of the ability to predict ice thickness to the stakeholders of the shipping industry operating in such zones, and in particular, in light of regulatory requirements and safety measures.

This study seeks to leverage machine learning (ML) and artificial intelligence (AI) technologies to accurately predict sea ice thickness in the Arctic, with a specific focus on the NSR. Unlike most existing research, which predominantly concentrates on predicting sea ice concentration, this study aims to address the significant gap in predicting sea ice thickness—a critical factor for maritime activities in the Arctic. Notably, to the best of the authors' knowledge, no previous work has applied Long Short-Term Memory (LSTM) models for predicting sea ice thickness along the NSR and its surrounding regions, nor has there been comprehensive coverage of huge Arctic expanses with a high level of accuracy in this context. By

Table 1 Comparison of LSTM with other regression methods

Method	Description	Strengths	Weaknesses	Comparison with LSTM
ARIMA	AutoRegressive Integrated Moving Average	Well-suited for univariate series; interpretable	Limited to linear relationships; data must be stationary	LSTM can handle non-linear, non-stationary data
ETS	Exponential Smoothing: Weighted averages of past observations	Simple to implement; good for short-term forecasting	Struggles with complex seasonality; less adaptive to new data	LSTM adapts to complex patterns and seasonality more flexibly
SARIMA	Seasonal ARIMA (extension of ARIMA)	Good for capturing seasonal effects	Requires manual parameter tuning; struggles with non-linearities	LSTM handles seasonality and trends without explicit tuning
VAR	Vector AutoRegression for multivariate series	Effective for multivariate series with dependencies	Assumes linearity; complex parameter estimation	LSTM handles multivariate data with non-linear dependencies
Prophet	Decomposition-based forecasting model	Easy to use; handles holiday effects and outliers well	Limited handling of complex non-linear relationships	LSTM can learn intricate non-linear dependencies directly
LSTM (Long Short-Term Memory)	Neural network capable of learning sequences and dependencies	Captures complex non-linear relationships; can handle long-term dependencies	Computationally intensive; requires more data and tuning; less interpretable	LSTM excels in capturing complex, non-linear, and sequential dependencies, unlike most traditional methods

utilizing LSTM models fed with data from the Copernicus database (Copernicus, 2020), this study aims to develop a predictive model that can identify safer and more efficient maritime shipping routes along the NSR, contributing to enhanced operational planning and adherence to regulatory safety measures in these challenging environments.

The remainder of this paper is structured as follows. Section 2 develops the methodology used, including data collection and processing and the development of the LSTM model. Section 3 details the obtained results of ice thickness prediction, while Sect. 4 provides concluding remarks.

2 Methodology

Figure 2 provides the framework diagram of the study. The study starts with data definition and collection, then data preprocessing, followed by the development of the LSTM model,

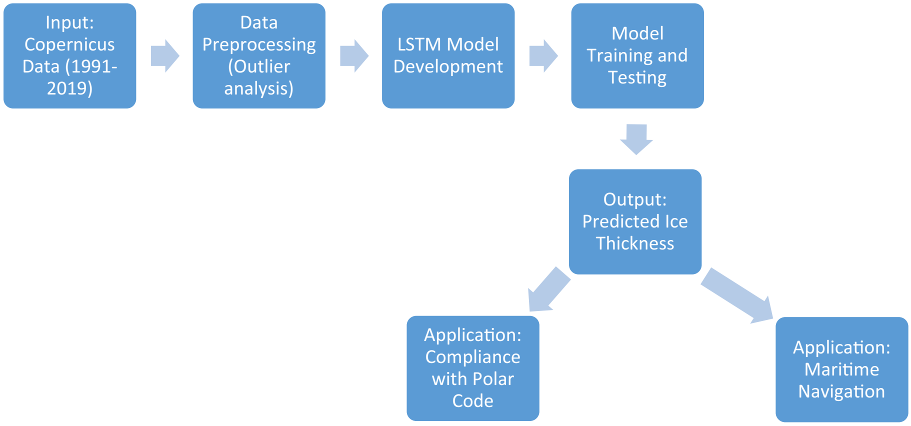


Fig. 2 Framework diagram of the study

and the training and testing phases, so that the predicted values of ice thickness are obtained and can be used for compliance and maritime navigation safety purposes. These steps are detailed in the following sections.

2.1 Area selection

While the ice melts, the density of navigation increases. However, ice melting does not mean that the ice completely disappears and, therefore, it still represents a direct risk for navigation, hence, anticipating this risk is paramount for safer navigation. Besides, as ice thickness within the NSR is not homogenous the conditions may highly differ from one area to another. Additionally, being able to provide a steady transit time is a key element for the transit of cargo, and as explained by Cariou et al (2019), ice impacts directly this parameter. Consequently, being able to anticipate ice thickness is a critical question for shipowners.

To develop a prediction model to foresee the ice thickness values in the Arctic, gathering ice thickness information over the past years is vital to train the LSTM model. The Arctic is huge, with an area of 5.5 million square miles (14.2 million square kilometers). With the computation power available and time restrictions, it is challenging to perform an analysis on the whole Arctic area. Specific areas of interest were therefore chosen. To select the areas, a search of the most common routes taken by vessels in recent years was conducted. Figure 3 from the NSR Information Office (2020) shows the voyages that took place in 2017. Similar data was also obtained for 2019.

From Fig. 3 it can be noticed that there are specific areas that are commonly used by vessels. In this study, based on the traffic data, four areas were selected to be used with the LSTM model for ice thickness prediction and to validate whether the used model was able to perform well and provide good results.

The selection of the four areas was mainly based on the commonly used routes for vessels, while also considering different locations such as open areas, areas between islands, and nearshore locations. The areas were selected using different geographical locations as follows:

- "Area 1" is between 22, 60 longitudes, and 68, 74 latitudes.
- "Area 2" is between 45, 80 longitudes, and 78, 90 latitudes.
- "Area 3" is between 90, 113 longitudes, and 77, 82 latitudes.

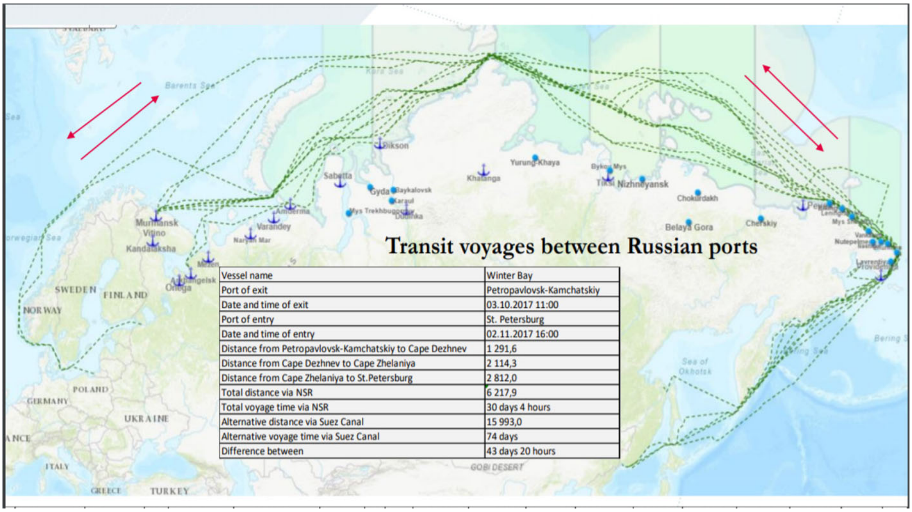


Fig. 3 NSR Shipping Traffic in 2017 (NSR Information Office, 2020)

- "Area 4" is between 117, 135 longitudes, and 74, 80 latitudes.

As can be seen in Fig. 4, the four Arctic areas included in this study cover a large part of the shores and navigational areas of the Arctic in addition to a remote area (Area 2) that is the northernmost. These areas cover a large part of the NSR including, the eastern, western, southern, and northern parts, and were selected for this reason.

Each area was considered fully in the analysis using its average ice thickness (per day or per month) that were calculated based on all its squares. In addition, each area was divided into subareas and these subareas were analyzed separately using their individual average ice thickness. More specifically, Area 1 was divided into 9 subareas, Areas 2 and 3 into 4 subareas each, and Area 4 into two subareas, which represents a total of 19 analyzed subareas. This

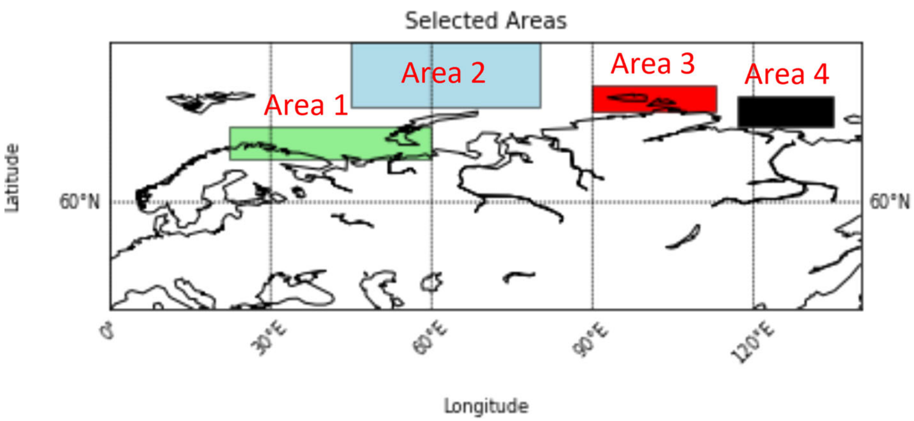


Fig. 4 The four Arctic areas considered in the study

approach provides a strong potential of generalizability of the results to the whole Arctic area.

2.2 Data collection

In this study, we obtained the required data from the Copernicus database (Copernicus, 2020) for the period from 1-1-1991 to 31-12-2019. Copernicus is the European Union's earth observation and monitoring program. It uses different technologies embedded on satellites in space to measure systems on the ground, in the sea, and in the air to deliver data and information that is open and free upon registration. The chosen dataset is ARCTIC_REANALYSIS_PHY_002_003 (Copernicus, 2020). The spatial extent of the dataset covers the Arctic Ocean between Latitudes 50°–90° and Longitudes − 180° to 179.88°. Its spatial resolution is 12.5 × 12.5 km. The current temporal extent of the dataset is from 1 Jan 1991 to 31 Dec 2023 and is available under the name ARCTIC_MULTIYEAR_PHY_002_003 (Copernicus, 2024). The data contains multiple variables as follows: zonal velocity (u), meridional velocity (v), sea surface height (ssh), temperature, salinity, sea ice concentration (fice), sea ice thickness (hice), sea ice zonal velocity (uice), sea ice meridional velocity (vice), snow thickness (hsnow), ocean barotropic stream function (bsfd), ocean mixed layer thickness (mlp), latitude, and longitude. The ice thickness product is the weekly merged CS2SMOS dataset from Alfed Wegener Institut (Ricker et al., 2017), combining thin sea ice measurements from the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) mission with the thick sea ice retrievals from another ESA mission, namely CryoSAT2. The merged product is assembled on a weekly basis and contains mapping errors representing uncertainty estimates. An additional uncertainty was therefore added to account for retrieval errors increasing linearly as a function of thickness. In the current dataset, the data is available in three different temporal resolutions: daily, monthly, and yearly average fields and a spatial resolution based on a grid of squares sided 12.5 km (Copernicus, 2024). The dataset is updated biannually. The format of the downloaded data is Unidata's Network Common Data Form (NetCDF). It is possible to download the data using different download options of subsetter and file transfer protocol (FTP). In the subsetter option, the user can filter the data using the following criteria: geographical area, depth, time range, and variables. The subsetter option enables manual downloads and downloads using a programming script. The second option is FTP, where one can connect to the FTP server with Copernicus Marine Service credentials to select dataset files. However, this option does not allow for any filtering criteria where one has to download the data as they are. For detailed information regarding the dataset and the download instructions, we refer the reader to Copernicus Marine Services (2024).

2.3 Data preprocessing

Before using the data, it is very important to make sure that it is organized and in the correct shape to be fed into the LSTM model. The data downloaded cannot be used directly as it is in the special NetCDF format. The general structure of NetCDF files is shown in Fig. 5.

Figure 5 shows that the data structure has multi-dimensional variables stored in one place. However, the LSTM model accepts the data in a tabular form of rows and columns containing each variable in a separate column, as shown in Fig. 6.

It was therefore necessary to convert the multi-dimensional variables of the NetCDF file into a tabular form. To achieve that, the following steps were taken:

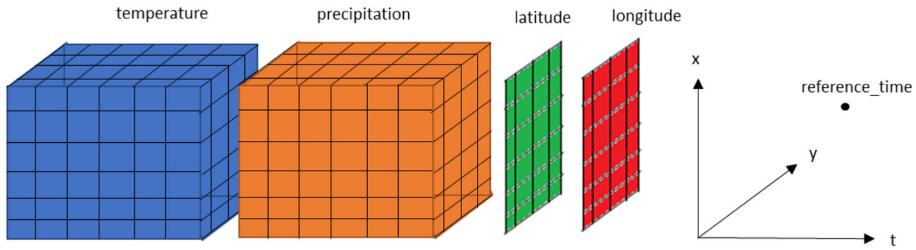


Fig. 5 NetCDF file structure (Hoyer & Hamman, 2017)

	A	B	C	D	E	F	G
1	uice	bsfd	salinity	vice	hsnow	ssh	fice
2	0	-1831334	34.93087	0	-6.66E-16	-0.32205	8.33E-16
3	0	-1831334	34.90271	0	-6.66E-16	-0.31954	8.33E-16
4	0	-1831334	34.87661	0	-6.66E-16	-0.31733	8.33E-16
5	0	-1526112	34.87181	0	-6.66E-16	-0.31616	8.33E-16
6	0	-1526112	34.87455	0	-6.66E-16	-0.31562	8.33E-16
7	0	-1526112	34.87661	0	-6.66E-16	-0.31471	8.33E-16
8	0	-1526112	34.87524	0	-6.66E-16	-0.31272	8.33E-16
9	0	-1526112	34.87112	0	-6.66E-16	-0.30947	8.33E-16
10	0	-1526112	34.86288	0	-6.66E-16	-0.30603	8.33E-16
11	0	-1526112	34.84296	0	-6.66E-16	-0.30316	8.33E-16
12	0	-1526112	34.81893	0	-6.66E-16	-0.30138	8.33E-16
13	0	-1526112	34.80725	0	-6.66E-16	-0.3012	8.33E-16
14	0	-1220889	34.82373	0	-6.66E-16	-0.30183	8.33E-16
15	0	-1220889	34.85052	0	-6.66E-16	-0.29831	8.33E-16
16	0	-1220889	34.83747	0	-6.66E-16	-0.28737	8.33E-16
17	0	-915667	34.79352	0	-6.66E-16	-0.27005	8.33E-16
18	0	-610445	34.72965	0	-6.66E-16	-0.24754	8.33E-16
19	0	-610445	34.65823	0	-6.66E-16	-0.22175	8.33E-16
20	0	-305222	34.60191	0	-6.66E-16	-0.199	8.33E-16
21	0	0	34.56758	0	-6.66E-16	-0.18399	8.33E-16
22	0	0	34.55247	0	-6.66E-16	-0.17765	8.33E-16
23	0	0	34.55109	0	-6.66E-16	-0.17797	8.33E-16
24	0	0	34.55041	0	-6.66E-16	-0.18043	8.33E-16
25	0	0	34.5353	0	-6.66E-16	-0.17677	8.33E-16
26	0	0	34.47555	0	-6.66E-16	-0.16639	8.33E-16
27	0	0	34.41237	0	-6.66E-16	-0.15529	8.33E-16
28	0	305222.4	34.37117	0	-6.66E-16	-0.14743	8.33E-16
29	0	-1526112	34.93911	0	-6.66E-16	-0.32181	8.33E-16

Fig. 6 Final MS Excel sheet

1. Take the first row of the data and transpose it, then save it into a new variable.
2. Take the transpose of the following row and append it to the first one.
3. Repeat 1 and 2 until all the rows are transposed and appended to the new variable.
4. Repeat 1, 2, and 3 for all the variables in the downloaded NetCDF file.

The result of the previous steps is a separate variable for each parameter mentioned earlier in Sect. 2.2 that is exported to an MS Excel file. The downloaded data does not have a separate variable for the year, month, and day, and such a temporal resolution can be generated manually if needed.

To complete this data preprocessing step and prepare the data files for the LSTM model, Python programming language and Jupyter Notebook integrated development environment were utilized. More specifically, Python modules and libraries such as Tensorflow, Keras, Pandas, Numpy Matplotlib, and Sklearn were used.

2.4 Outlier analysis

This section discusses the analysis performed on the downloaded data to check for any outliers. Outliers are data points that are far away from the rest of the data and are usually considered abnormal values. It is believed that having many outliers in the data affects the prediction process.

To identify outliers, a boxplot is performed on different areas' data to check whether the data contains any outliers. Boxplot is a type of chart used to show the distribution and skewness of the data (McLeod, 2019). The boxplot results for Area 2 and Area 4 are shown in Figs. 7 and 8.

Figure 7 shows the resultant boxplot performed on Area 2 and its subareas. One can notice that there are no data points that fall outside the whiskers or the boxplot, which means that there are no outliers in the data.

Figure 8 shows the boxplot of Area 4 and its subareas. It can be noticed that there are some outliers. However, the number of outlier points found is very low compared to the amount of data available. Also, outliers are considered incorrect readings that sometimes appear in the data due to errors in the data collection or a fault in the used tool, but in our case, these outliers are likely to be real data points that appeared due to the large variability and the continuous change of the ice thickness values across different periods of the year. In addition, when dealing with ice thickness values that are used as a basis to choose the type of vessels to use for Arctic sailing and the value of the sailing speed, it would be useful to include these extreme values in the training process of the model. Hence, it was decided to keep the outliers in the data and not to remove them. Similar results were obtained for Areas 1 and 3.

The decision to keep outliers was made because of the unexpected extreme cold and warm weather that the Arctic is facing. Due to the global warming phenomena, there have been instances where predicted extreme cold turned out to be clear from ice and vice versa. Indeed, despite the available predictions and the expectation of the Arctic with no ice, the ice conditions in 2013 were rather difficult, countering what was believed. Humpert (2014) mentioned that the shipping is subject to the intra-seasonal variability of ice conditions, which is a key economic obstacle of the NSR. Humpert (2014) pointed out that in 2013, difficult ice conditions between late August and mid-September had a noticeable impact on shipping activity. Furthermore, inter-annual variations of ice conditions will also remain significant. While 51 vessels had transited the NSR by October 1 in 2013, no transits had been recorded by the same date in 2014.

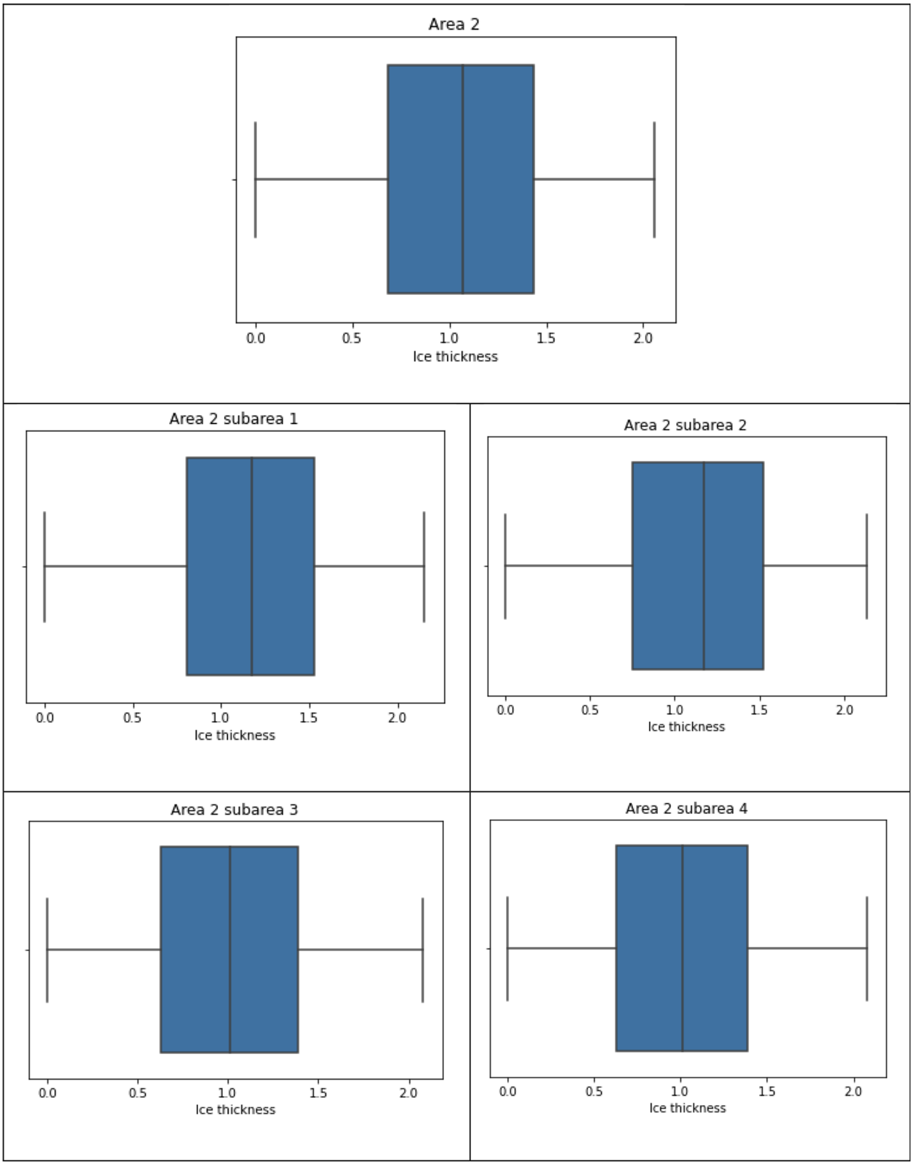


Fig. 7 Area 2 outlier analysis using boxplot

However, to check the impact of removing outliers on the results, we performed further analysis. Indeed, we removed the outliers and compared the results with those with outliers and found no significant difference. The results are shown in Sect. 3.4.

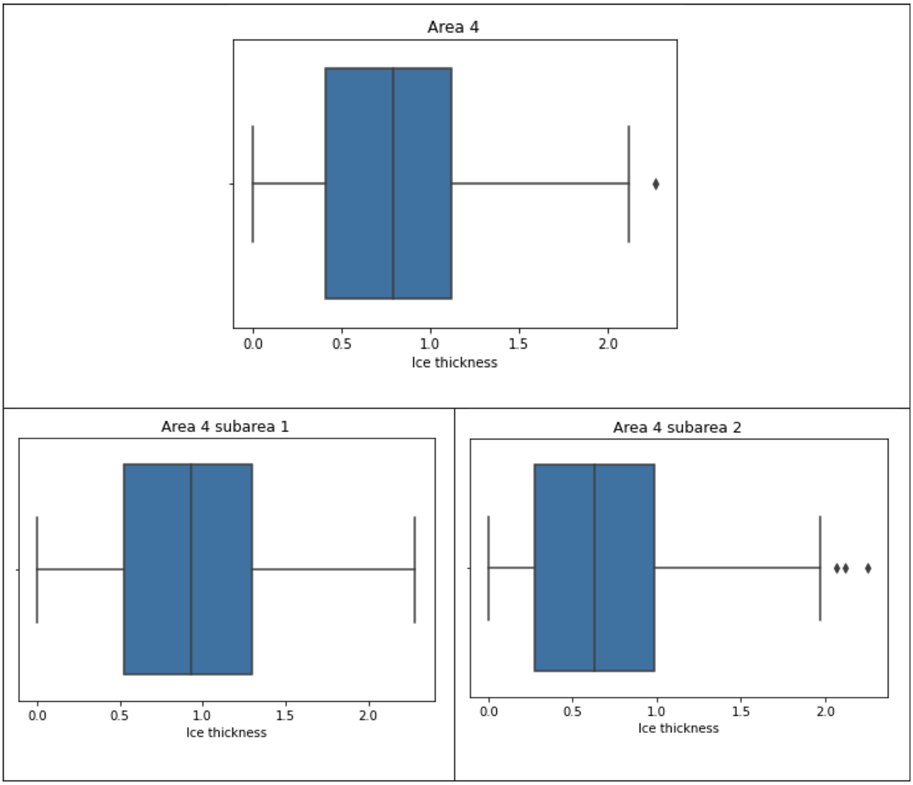


Fig. 8 Area 4 outlier analysis using boxplot

2.5 Design of experiments

In developing the LSTM model, the first step is deciding on which input data to use in the model. A correlation analysis was performed between ice thickness and the variables available inside the dataset mentioned in Sect. 2.2, and the results are reported in Table 2.

Table 2 shows a pairwise correlation analysis between sea ice thickness and the rest of the variables available in the dataset. The closer the value is to 1, the stronger the positive relationship between the variables. The closer the value is to − 1, the stronger the negative relationship. The values in Table 2 illustrate that there is a minimal correlation between the ice thickness and the other variables except for sea ice concentration (fice), and snow thickness (hsnow). However, the correlation values are not very close to 1 or −1. Therefore, it was decided to use the value of ice thickness only as an input for the LSTM model.

Table 2 Correlation analysis between sea ice thickness and other variables

uice	bsfd	ualinity	hsnow	ssh	fice	mlp	u	v	Temperature
−0.0125	0.1211	−0.330	0.735	0.026	0.760	−0.132	−0.0324	−0.082	−0.424

A series of experiments were conducted, as will be detailed in the Results section (See Sect. 3), on the input data and the model's parameters to check the effect of these changes on the model output. The series of experiments are listed as follows:

- Using daily ice thickness values as an input of the model versus using monthly ice thickness values and their effect on the model prediction.
- Applying the model to different geographical areas.
- Testing different percentages to split the data between training data and testing data.
- Reducing the amount of data considered as input.
- Changing the prediction period.
- Checking the parameters variability and their correlation with the quality of the output.

As previously explained, the RIO (Risk Index Outcome) defined in POLARIS, enables decision-makers to decide on the capacity of vessels to sail in the selected area and determines the optimal sailing speed enabling sailing in safe conditions. As emphasized by Fedi et al. (2020) and Browne et al., 2022, POLARIS is a decision-making tool that makes navigation safer.

POLARIS translates the ice thickness values into Risk Index Values (RIV) as shown in Table 3[2] depending on the ice thickness and the type of ice. In addition, the obtained $\boldsymbol{RIV_i}$ values of all the ice types $\boldsymbol{i} = 1, \ldots, \boldsymbol{n}$ and their corresponding concentrations ($\mathrm{C_i}$) are then combined to calculate the RIO value using:

$$\mathrm{RIO} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{RIV_i}.\mathrm{C_i} \tag{1}$$

With n = 11 being the total number of ice categories. Moreover, the obtained RIO value is between -8 and + 3.

As stressed by Marchenko (2014), speed is a key factor in terms of risk management. An inappropriate speed can lead to a marine incident or marine casualties (IMO, 2008). Here again, POLARIS provides some information regarding the suitable speed based in the RIO depending on the ice thickness that vessels may encounter as shown in Table 4.

As shown in Table 4, for instance, if the RIO value is positive for a given vessel, then the vessel can sail at its optimal speed in the area under consideration.

2.6 Developing the LSTM model

ANN is one of the machine learning types that can be used in classification, pattern recognition, clustering, and prediction in many fields and disciplines (Abiodun et al., 2018). It is a computational approach used to analyze data and solve specific problems inspired by the human nervous system (Aichouri et al., 2015). A basic ANN consists of an input layer, a hidden layer, and an output layer, each of which includes different neurons connected with each other with weights, and each of which has an activation function. The number of neurons in each layer differs from one application to another.

Moreover, LSTMs are a special kind of Recurrent Neural Networks (RNNs), capable of learning long-term dependencies and patterns, and were therefore used in this work to predict ice thickness values. RNNs differ from other types of ANNs as they have loops and feedback connections, unlike many other ANNs that use the feedforward standard strategy.

LSTM contains units called "memory blocks" and considers the information learned in a short period (short-term memory), or the hidden state $\boldsymbol{h_t}$, as shown in Fig. 9, and uses it

[2] https://climate.copernicus.eu/climate-indicators/sea-ice

Table 3 Risk index values

Vessel category	Vessel class	Ice free	New ice	Grey ice	Grey white ice	Thin 1st year ice	Thin 2nd year ice	Medium 1st year ice	Thick 1st year ice	2nd year ice	Light muti-year ice	Heavy multi-year ice
Polar Class (Cat. A)	PC-1	3	3	3	3	2	2	2	2	2	1	1
	PC-2	3	3	3	3	2	2	2	2	1	1	0
	PC-3	3	3	3	3	2	2	2	2	1	0	− 1
	PC-4	3	3	3	3	2	2	2	1	0	− 1	− 2
	PC-5	3	3	3	3	2	2	1	1	− 1	− 2	− 3
Polar Class (Cat. B)	PC-6	3	2	2	2	2	1	1	0	− 2	− 3	− 3
	PC-7	3	2	2	2	1	1	0	− 1	− 3	− 3	− 4
Ice class	1AS	3	2	2	2	2	1	0	− 2	− 3	− 4	− 5
	1A	3	2	2	2	1	0	− 1	− 3	− 4	− 5	− 6
	1B	3	2	2	1	0	− 1	− 2	− 4	− 5	− 6	− 7
	1C	3	2	1	0	− 1	− 2	− 3	− 5	− 6	− 7	− 8
Category II	Not ice class	3	1	0	− 1	− 2	− 3	− 4	− 6	− 7	− 8	− 8

Table 4 POLARIS decision rules based on RIO values. *Source* IMO (2016)

RIO_{SHIP}	Ice-classes PC1-PC7	Ice-classes below PC7 and ships not assigned an ice-class
RIO $\geq$ 0	Normal operations	Normal operations
$-$ 10 $\leq$ RIO $<$ 0	Elevated operational risk	Operations subject to special consideration
RIO $\leq$ $-$ 10	Operations subject to special consideration	Operations subject to special consideration

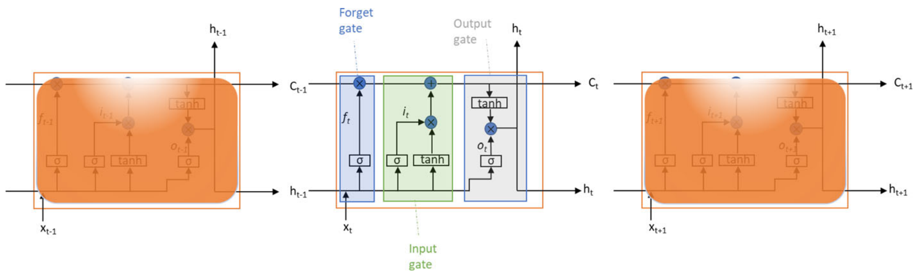


Fig. 9 LSTM model

for training the information stored in the long-term memory, or the memory cell state $\boldsymbol{C_t}$ as shown in Fig. 9. Each memory block consists of three gates, a forget gate, $\boldsymbol{f_t}$, an input gate $\boldsymbol{i_t}$, and an output gate, $\boldsymbol{o_t}$, to control the flow of information (Kırbaş et al., 2020).

Indeed, the forget gate enables information to be discarded from the long-term memory by calculating the percentage of the long-term memory to pass to the next stage by using a sigmoid activation function (σ). The input gate adds information from the current timestamp (short-term memory) to the current long-term memory as it is deemed useful by using a sigmoid and a tanh activation functions. Finally, the output gate provides the hidden state of the cell at time t, using a sigmoid and a tanh activation functions, based on a combination of the updated memory, the new information collected ($\boldsymbol{x_t}$), and the previous short-term memory and therefore allows to calculate the new short-term memory.

The initial plan was to consider all the variables in the dataset as inputs to the LSTM to predict the value of ice thickness. However, multiple challenges were faced. The first challenge was the large amount of data that was fed to the ANN. Since the data is recorded day by day between the period of 1-1-1991 and 31-12-2019 and it is for a wide geographical area, we ended up with 10,592 days and 8514 locations which led to 90,180,288 values for each parameter. This large amount of data required huge processing power and time to put it in a format that can be used.

Throughout the study, multiple LSTM models were developed to examine which architecture would be able to predict the ice thickness with the lowest possible error. The developed models were split into two categories. The first model architecture was designed for the daily ice thickness data as an input to examine how good the daily data was at predicting mid to long-term ice thickness values. The second model architecture was modified to take monthly input data, and it was compared subsequently with the model with daily input data.

In developing any LSTM model, some parameters should be specified, such as the number of layers, neurons, batch size, epoch, activation function, and optimizer. In the following paragraphs, some of these parameters are discussed.

Moreover, in order to train the developed model, an optimizer should be used. Optimizers are sets of methods and algorithms that increase the accuracy and reduce the error of the prediction by changing the weights and learning rate of the ANN (Bushaev, 2018). In this study, Adam optimization algorithm was used. It is widely used in deep learning models and represents an extension of stochastic gradient descent, which has been widely used in deep learning applications in recent years. The optimization method is used to update network weights iteratively based on training data. Adam optimization algorithm is straightforward to use, uses little memory, is appropriate for problems with a large amount of data and parameters, and suits non-stationary objects and problems with noise. Adam is a widely used ANN algorithm mainly due to its fast speed and high-quality results (Brownlee, 2017; Bushaev, 2020). In order to use the optimizer, two parameters need to be fixed, namely, the batch size and the number of epochs and they differ from one model to another (Murphy, 2017). Indeed, the batch size is a hyperparameter that defines the number of samples to work through before updating the internal model parameters. In addition, the epoch is a hyperparameter that defines the number of times that the learning algorithm will work through the entire training dataset. In the following passages, some of the specific values of the parameters are discussed in more detail.

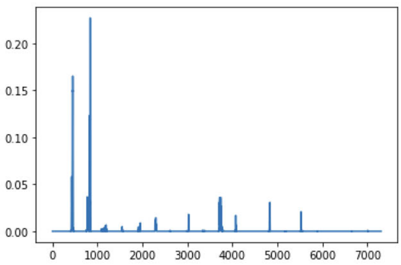
According to the literature (Abhinav et al., 2017; Cabaneros et al., 2017; de Carvalho Paulino et al., 2018; Gholami et al., 2015; Marín-Blázquez & Schulenburg, 2006; Niedbała, 2019), ANNs best perform when 70–80% of the data is considered as training, and the remaining 30–20% is considered as testing.

The architecture of the LSTM model with daily input data used 50 epochs and a batch size of 32. Moreover, the used activation function was the Rectified Linear Unit (ReLU) function, although other activation functions were tried but resulted in lower quality predictions, the optimizer 'Adam', the loss function was the mean square error (MSE) and the split between training and testing was 70–30% and 90–10%. For the monthly data model, the same parameters were used except the number of epochs was set to 200. As will be detailed later, the setting with 90–10% resulted in better results which may be due to overfitting.

As mentioned earlier, the LSTM neural network stores information from the past and uses it to predict values in the future. In developing the LSTM model, the number of values used to predict future values should be specified. In our model, we opted for the model to save information covering 2555 past days to predict 30 future days. Once all the parameters are set, the data is normalized and transformed into values between zero and one, split into two parts: the oldest 70% are used for training, and the most recent 30% for testing. One area out of the four areas mentioned earlier was chosen for this model (Area 1).

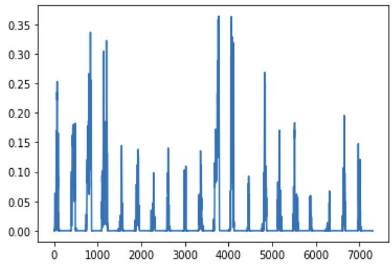
The first step taken in developing the LSTM model with daily data as input was to analyze the behavior of the ice thickness across different locations within the same area. This was done by plotting the ice thickness values for all the locations inside the area. By analyzing the behavior of the ice thickness in the area, one can notice that there is a high variability as shown in Fig. 10, which shows the ice thickness values in meters of different locations of Area 1 in the period between 2000 and 2019 in days.

Figure 10 shows the ice thickness values in meters of different locations in the period between 2000 and 2019 in days. Two observations can be drawn: the cyclical (seasonal) nature of ice thickness, which corresponds to the seasons of the year, and the variability of ice thickness among the same seasons of the considered years (period).



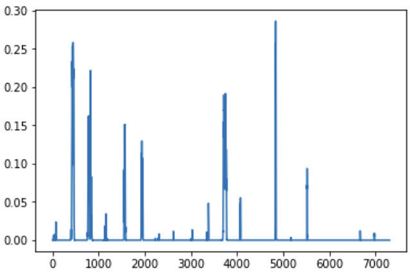
Latitude: 68.30242

Longitude: 40.601295



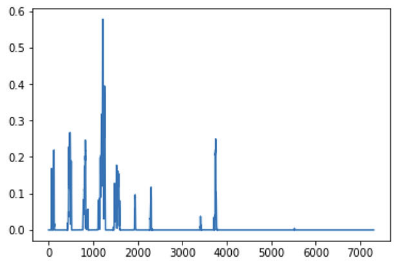
Latitude: 68.45056

Longitude: 42.343853



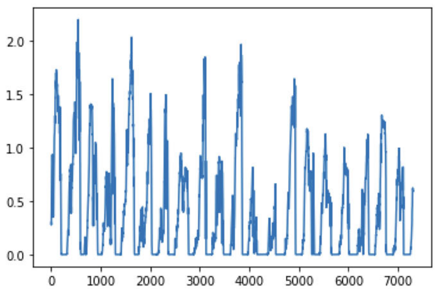
Latitude: 69.66757

Longitude: 45



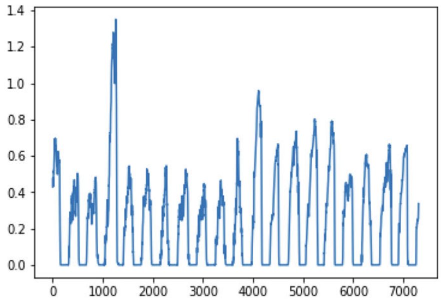
Latitude: 74.410

Longitude: 50.751404



Latitude: 73.1547

Longitude: 58.3924



Latitude: 69.5753

Longitude: 59.4897

Fig. 10 Ice thickness values in Subarea 3 of Area 1

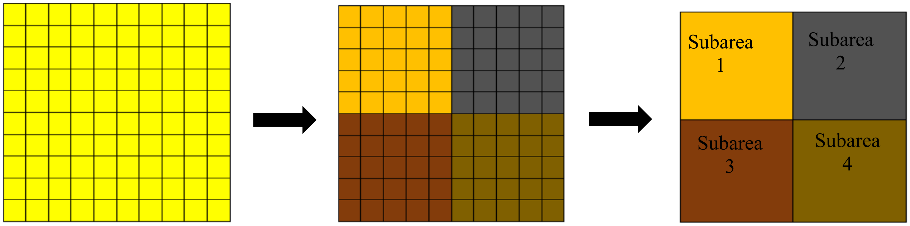


Fig. 11 Area splitting

To start the analysis, an area with changing ice thickness values was chosen to run the LSTM model. The chosen location latitude and longitude are 73.1547 and 58.3924, respectively. The prediction results will be discussed later.

For the LSTM model with monthly input data, 200 epochs and a batch size of 32 were used. For this model, the input dataset considered for the LSTM model is the monthly collected data between the period of 15-1-1991 and 15-12-2019, where each value is the mean ice thickness value of the month at the time of collection. The number of future values to be predicted in the future should be specified. In this study, we chose the model to use historical information treating 88 months and to predict eight future months. Once all the parameters are set, the data is normalized and transformed into values between zero and one, split into 70% for training and 30% for testing, and then fed to the model. The area under consideration for this model is Area 1. In this model, a different approach was taken into consideration, which is splitting the full area into smaller subareas. This way, the variability of ice thickness values between far locations will not affect the prediction process. The split is done by taking nearby locations and averaging their ice thickness values, resulting in a subarea, as shown in Fig. 11. This process resulted in nine subareas shown in Fig. 12.

Figure 11 demonstrates the method used in splitting the areas into subareas. Each area is split into an arbitrary number of subareas. The split is done by taking the nearby locations next to each other and averaging their ice thickness values, resulting in one mean value of all the nearby locations. Repeating the steps for all periods will finally result in the data of the subarea.

Figure 12 shows the final resultant subareas for Area 1 (sorted from Subarea 1 through 9).

3 Results

Our results are based on different layers of analysis. First, we demonstrate the capacity of the LSTM model to predict ice thickness for a short-term period based on daily input data, stress its limits, and use the model with monthly input data. Second, we change both the share of training data compared to the current habits and the amount of historical input data used. Third, we consider two different lengths of the prediction period and discuss the obtained results. Fourth, we focus on the input parameter variability and their effect on the accuracy of the predictions.

It is worth pointing out that in all the results reported hereafter, the chosen performance criterion to judge the model output was the Mean Square Error (MSE). As mentioned earlier, we used the optimization function “Adam”. “Adam” is based on a gradient descent optimization algorithm that does not commonly yield good results from one run (Cheridito et al., 2021). Therefore, the shown results are taken after running the model 10–20 times and

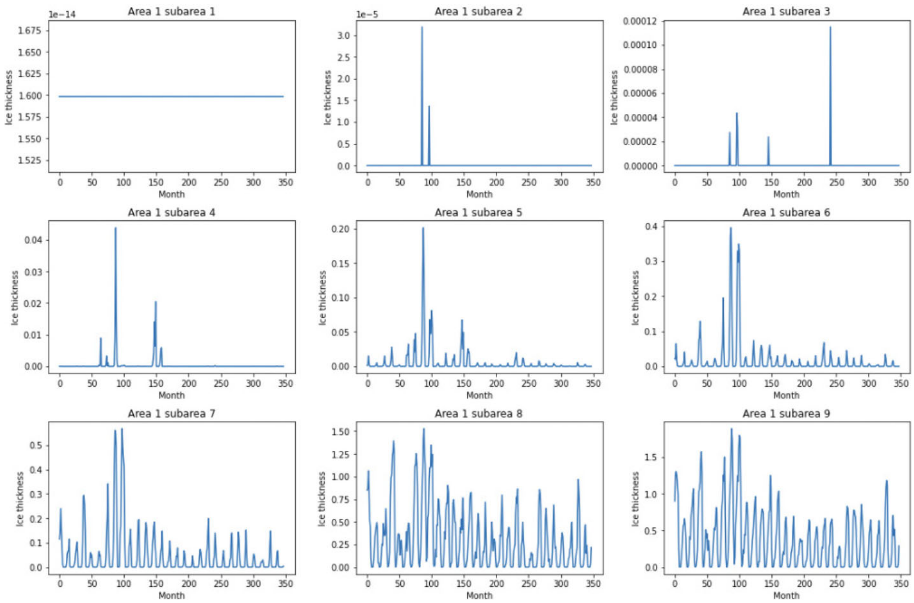


Fig. 12 Ice thickness values in the subareas of Area 1

then considering the runs with the lowest error results following the approach proposed by Yibre and Koçer (2021). The weight for the best run result can be saved and used for future predictions.

All the figures hereafter show the ice thickness values in meters on the y-axis and the temporal resolution on the x-axis (days in the daily input model and months in the monthly input model).

3.1 Effect of the resolution of the input data and the length of the prediction period

It is agreed that ice thickness varies based on both the geographical location (area) and the calendar (day or month of the year). As Area 1 is the busiest area in the Russian Arctic and is subject to low ice thickness compared to the rest of the NSR, we applied our model to predict the ice thickness in Area 1 on a daily basis using the daily input data. Figure 13 highlights the prediction of thickness. As some negative values were obtained, which is impossible, we corrected the obtained results by applying a postprocessing function to set the negative values to zero. Figure 14 shows the obtained results after postprocessing was applied. Based on Figs. 13 and 14, it appears that the LSTM is a good prediction tool with an average MSE value of 0.0468. Notwithstanding, if these results are based on daily data, they offer a projection over 30 days.

If we increase the projection to 60 days for the same area (Area 1), then the MSE value becomes 0.051, which remains below 0.1, while increasing the length of the prediction period further to 120 days makes the prediction quality decrease with an MSE of 0.147. Thus, it appears that the daily data combined with the LSTM are limited in terms of prediction duration to a short period of nearly one month.

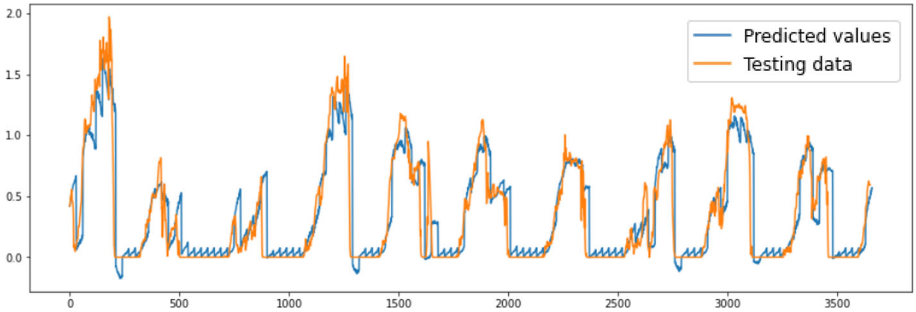


Fig. 13 Preprocessed predicted values versus Testing ice thickness values for Area 1

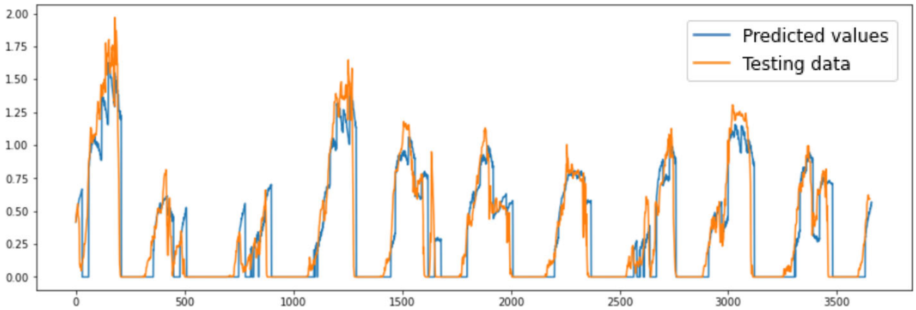


Fig. 14 Postprocessed predicted values versus testing ice thickness values for Area 1

To counteract this loss of reliability, we used monthly input data instead of daily ones. By using monthly values, we were able to reach an MSE of 0.00175, 0.0144, and 0.0479 in Subareas 7, 8, and 9 of Area 1, respectively. Hence, if stakeholders want to predict ice thickness for at least six months, it appears that monthly data is more accurate than daily ones.

To confirm the better accuracy of the LSTM model that is based on monthly input data, and since the ice thickness may change from one area to another, we decided to use the monthly data model for Areas 2, 3, and 4. Furthermore, ice thickness variability from one year to another represents one of the main challenges for forecasting ice thickness.

Based on data collected in those three areas, we obtained an MSE of 0.118, 0.093, and 0.068 for Area 2, Area 3, and Area 4, respectively, as shown in Fig. 15.

The subareas of the different areas under consideration were considered individually. For instance, Fig. 16 shows the ice thickness historical data variability in Area 2 and its two subareas. One can notice that there is a large variability in the ice thickness values throughout the whole period.

The MSE of the prediction of the subareas of Area 2 are above the limit of 0.1. The same results were obtained for most of the other subareas of Area 3 and Area 4. The difference in the MSE values can be explained by the existing level of variability of ice thickness input data from one zone to another. This variability can have a negative effect on the prediction.

Hence the results demonstrated that even if the monthly data seems to be more relevant than daily ones, when long-term projections are required, the variability of ice thickness still

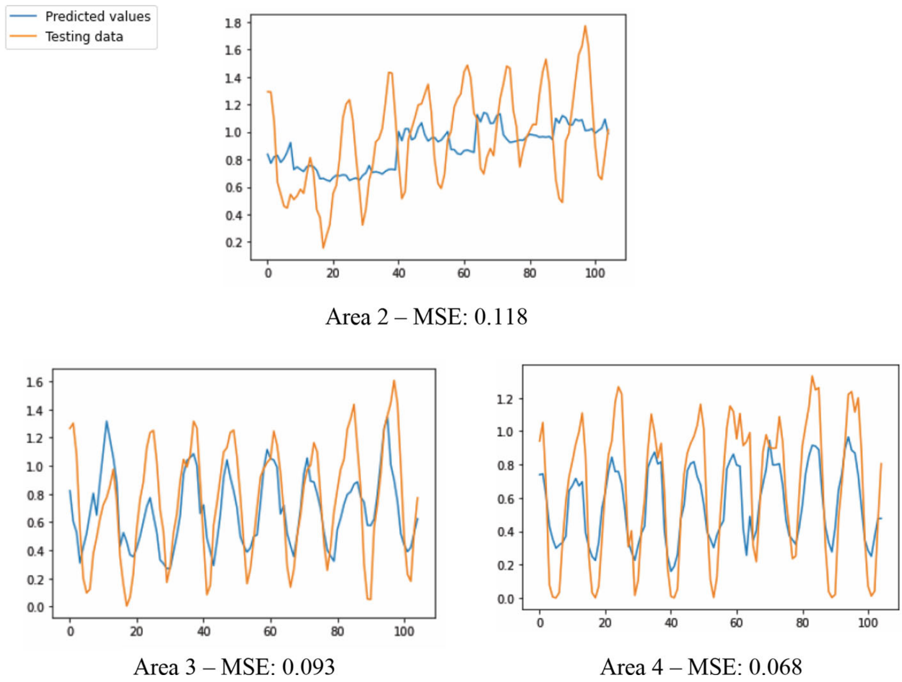


Fig. 15 Postprocessed predicted values versus testing ice thickness values for Areas 2, 3, and 4 using the monthly input data model

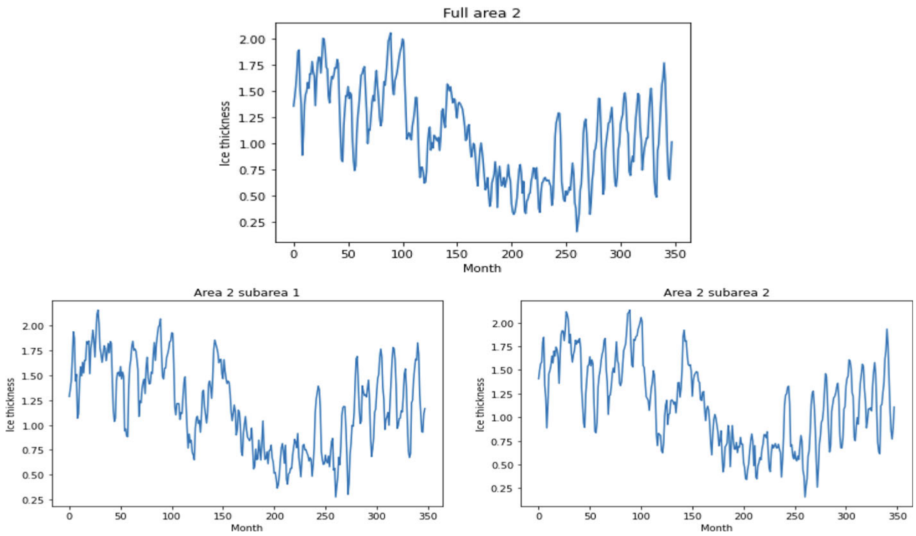


Fig. 16 Historical data of Area 2 and its subareas

has a negative impact on the prediction. These results call for more investigation into the reasons behind the low model performance in some areas compared to others.

Using aggregated data (monthly instead of daily), more training data (90% instead of 70%), and shorter prediction periods (one month instead of four months), improves the prediction accuracy. However, increasing the training set percentage should be weighted so that the risk of overfitting is avoided. Moreover, normalizing the data would potentially contribute to improving the accuracy of the model by controlling the variability. Another recommendation would be to cluster the zones based on the level of ice thickness and the temporal variability of ice thickness so that the model can be fine-tuned by cluster which is also thought to better improve the prediction.

3.2 Effect of the percentage of training data, amount of historical data, and length of the prediction period

As explained above, it is generally admitted that using 70% of the data for training is enough. Yet, due to the unpredictability of ice thickness, we considered changing this share to 90%. Figure 17 shows the comparison between the testing data and the predicted values for Areas 2, 3, and 4. By increasing the percentage of the training data, we reached an MSE value of 0.087 instead of 0.118, 0.053 instead of 0.093, and 0.048 instead of 0.068 for Area 2, Area 3, and Area 4, respectively. Hence, it appears that increasing the percentage of the training data enhances the level of reliability of the forecasting.

This part discusses the effect of considering the most recent data (years) on ice thickness as an input to the model, instead of taking the whole available dataset. We consequently used the data of the last 100 months for which ice thickness is available, instead of the 30 years of available data, and fed them to the LSTM model. We targeted estimating the ice thickness for the next eight months. We used the rule of 90–10% to split the used data into training and testing sets. The results of Areas 3 and 4 are shown in Fig. 18.

The obtained MSE values when both the rule of 90–10% and the most 100 months data are combined are 0.049 for Area 3 and 0.007 for Area 4 for eight months' estimation.

When considering a shorter forecast period of four months, the results are closer to the historical data (with the rule of 90–10%) for Area 2, but it is not the case for Areas 3 and 4. Indeed, for Area 2, the MSE dropped from 0.0276 to 0.0078. However, for Area 3, the MSE increased from 0.049 to 0.0653, and for Area 4, it increased from 0.007 to 0.0259.

Thus, reducing the length of the prediction period while increasing the percentage of the training data does not seem to be that efficient. Hence, reducing the number of forecasted months is only effective when more testing data is considered.

Reducing the available data from 30 years to 100 months did not result in a clear improvement in the prediction quality. Therefore, it is recommended to use as much as possible available data. This may be because LSTM learns from old data but also adjusts to the most recent data through long and short-term memories. This appears generally to be safer in better controlling the variability.

3.3 Variability of ice and prediction quality

This section discusses the analysis done on the parameters' variability and their correlation with the prediction quality. The model performed worse on the ice thickness data taken from Area 2 while having the best results with Area 4, as discussed in the previous sections. Initially, this was suggested to be a result of the ice thickness variability difference between

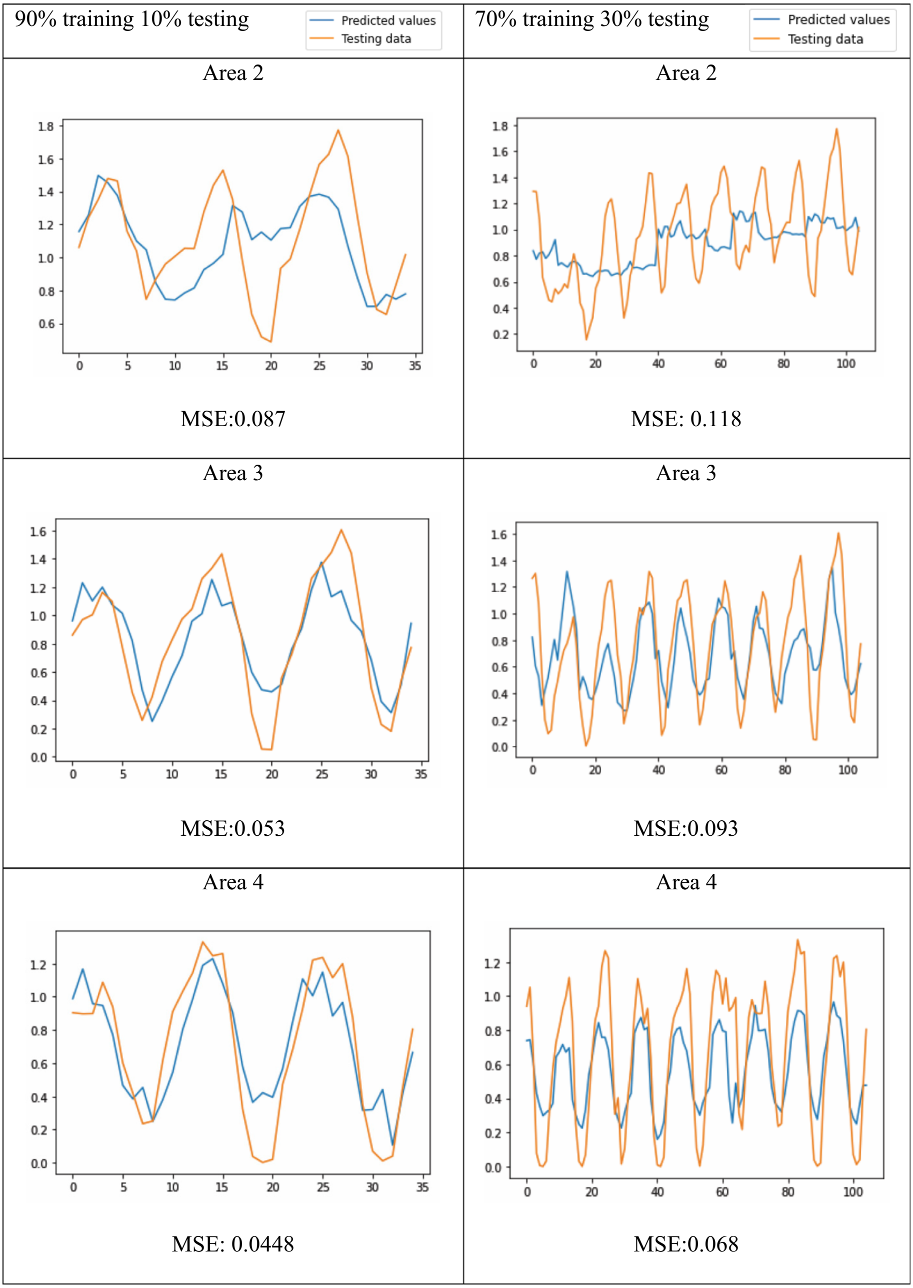


Fig. 17 The effect of changing training percentages on Areas 2, 3, and 4

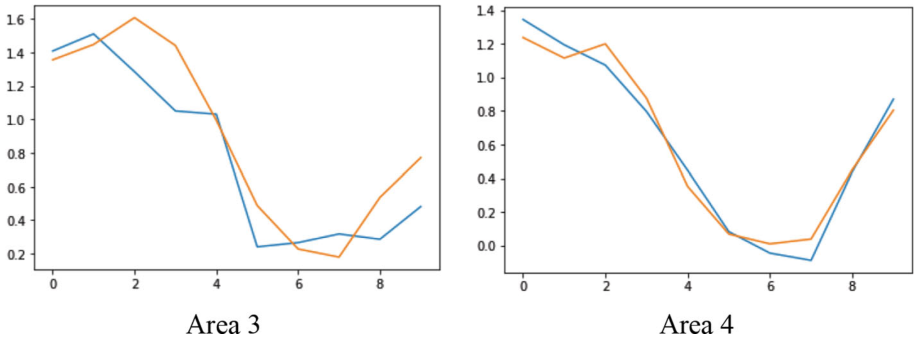


Fig. 18 Results based on considering the 100 most recent months as input to the LMST model

Table 5 Ice thickness standard deviation in Area 2 and Area 4

	90% training	10% testing	70% training	30% testing
Full area 2	***0.445***	***0.331***	***0.453***	***0.354***
Full area 4	0.484	0.432	0.496	0.412

In bold, we highlighted the smallest value among the different result in the column

Area 2 and Area 4. The standard deviation (std) of the ice thickness data in both areas is calculated and shown in Table 5.

Table 5 shows the standard deviation of the data in Area 2 and Area 4 for both training and testing periods. One can notice that the standard deviation in Area 2 across the training and testing datasets is lower than in Area 4. This finding suggests that the standard deviation of the data does not affect the model accuracy. To investigate more on this finding, the same analysis was done on two subareas of Area 2, namely Subarea 2 and Subarea 4. The results are shown in Table 6.

Table 6 shows the standard deviation of the data in Subarea 2 and Subarea 4 of Area 2 for both training and testing data. Indeed, the subarea with better results shown previously did have a lower standard deviation. This finding contradicts what was shown in Table 3. This suggests that the standard deviation by itself cannot be the only judging factor on whether the model will perform well or not. Moving forward, it was decided to investigate the impact of the amplitude of the ice thickness of each year, i.e., the difference between the maximum and the minimum ice thicknesses for each year, the impact of the minimum ice thickness of every year, and the impact of the maximum ice thickness of every year. In addition, the results indicate that there might be internal variability between the years that is not shown

Table 6 Standard deviation in Area 2—Subarea 2 and Area 2—Subarea 4

	90% training	10% testing	70% training	30% testing
Area 2—Subarea 2	0.462	***0.323***	0.469	0.371
Area 2—Subarea 4	***0.451***	0.381	***0.457***	***0.365***

In bold, we highlighted the smallest value among the different result in the column while calculating the standard deviation of the whole data. This variability is not the result of regular seasonality but of climate change that leads ice thickness to vary between the same month of different years. The seasonality can be identified generally by observing an increase in ice thickness values in winter and a decrease in summer. However, throughout the period under consideration (1991–2019), the results show that there is no clear trend in ice thickness for every season, which is due to the variability and the rapid change of ice thickness values. As a result, we considered the yearly values and showed them in Figs. 19, 20, and 21.

Figures 19, 20 and 21 show the values of the ice thickness amplitude, maximum, minimum, standard deviation, and mean for all the years under consideration. They were calculated to check whether there are internal variabilities across the years that did not show in calculating

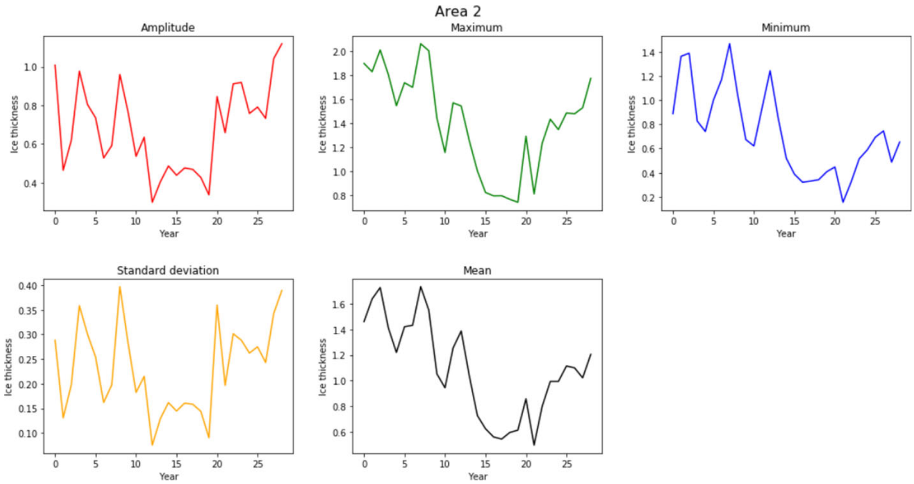


Fig. 19 Area 2 variability analysis

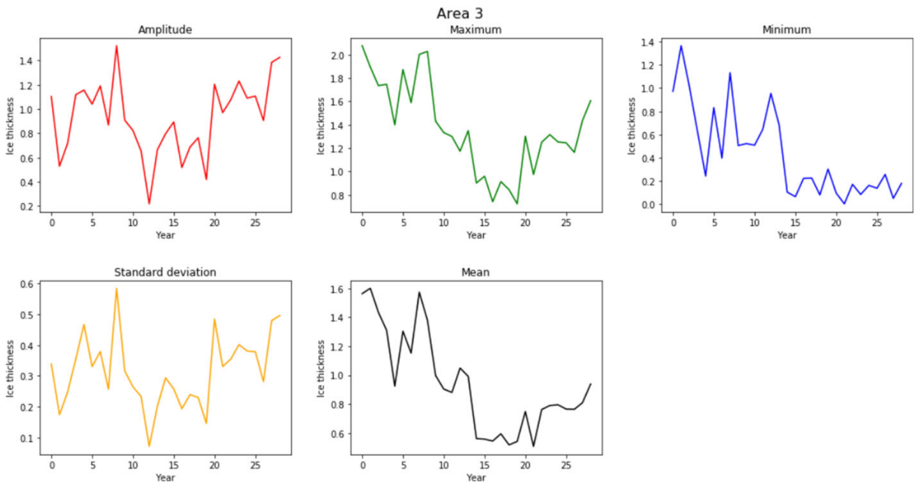


Fig. 20 Area 3 variability analysis

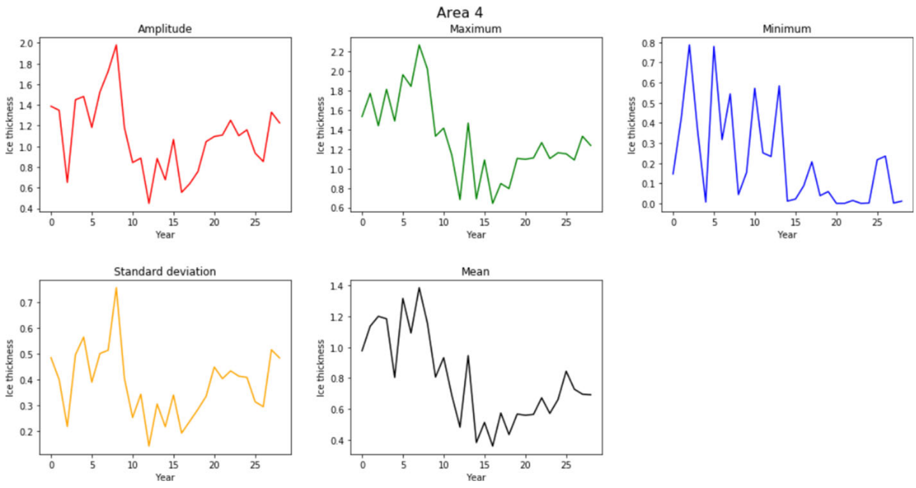


Fig. 21 Area 4 variability analysis

the standard deviation of the whole area. The focus of the analysis will be on the amplitude, maximum, and minimum values, for which the standard deviation and the mean were calculated using different periods of historical data as follows:

- "Second training period" includes the historical data of the period between months 11 and 22.
- "Whole training period" includes the months between 0 and 22.
- "Testing period" includes the data of the months between month 23 and month 29.

The results in Tables 7, 8, and 9 show which areas had lower standard deviation across different data periods. Looking at the data, it is hard to find a relationship between better prediction results and variability across the amplitude, maximum, and minimum values of the ice thickness. However, taking a closer look, one can notice that overall, Area 4 results are the best in terms of variability, where the difference between Area 4 variability and the lowest variability is low when it does not perform the best in terms of prediction quality. However, when Area 4 performs the best, the difference between the next close values is high, most of the time. To take this into the next step, that same analysis was conducted on the subareas for more investigation.

Table 10 shows that Subarea 2 of Area 2 has less variability in amplitude values across

Table 7 Mean and standard deviation of the amplitude values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2	0.501	***0.151***	0.612	***0.208***	0.896	***0.146***
Area 3	0.717	0.258	0.852	0.300	1.174	0.184
Area 4	0.833	0.222	1.086	0.400	1.122	0.173

In bold, we highlighted the smallest value among the different result in the column

Table 8 Mean and standard deviation of the maximum values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2	1.047	0.306	1.390	0.463	1.468	0.167
Area 3	1.042	***0.235***	1.376	***0.440***	1.323	0.149
Area 4	1.005	0.276	1.342	0.469	1.191	***0.089***

In bold, we highlighted the smallest value among the different result in the column

Table 9 Mean and standard deviation of the minimum values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2	0.546	0.313	0.778	0.391	0.572	0.144
Area 3	0.325	0.303	0.523	0.390	0.149	***0.067***
Area 4	0.171	***0.210***	0.255	***0.256***	0.068	0.107

In bold, we highlighted the smallest value among the different result in the column

Table 10 Mean and standard deviation of the amplitude values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2—Subarea 2	0.549	***0.150***	0.633	***0.185***	0.886	0.177
Area 3—Subarea 4	0.880	0.289	1.070	0.402	1.158	0.190
Area 4—Subarea 2	0.923	0.231	1.227	0.439	1.089	***0.070***

In bold, we highlighted the smallest value among the different result in the column

the training data and Subarea 2 of Area 4 has less variability in the testing data. The standard deviation and the mean of the maximum and minimum values across all defined periods are shown in Tables 11 and 12.

Tables 10, 11, and 12 show that the variability of Subarea 2 of Area 4 is much less across all periods for the amplitude, maximum, and minimum. From this observation, one can conclude that the prediction model developed will work better when the variabilities of the maximum ice thickness value and the minimum ice thickness value are low.

Table 11 Mean and standard deviation of the maximum values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2—Subarea 2	1.134	0.401	1.467	0.491	1.540	0.211
Area 3—Subarea 4	1.036	0.337	1.349	0.485	1.177	0.171
Area 4—Subarea 2	0.982	***0.275***	1.319	***0.481***	1.090	***0.071***

In bold, we highlighted the smallest value among the different result in the column

Table 12 Mean and standard deviation of the minimum values across Areas 2, 3, and 4

	Second training period mean	Second training period std	Whole training period mean	Whole training period std	Testing period mean	Testing period std
Area 2—Subarea 2	0.585	0.321	0.833	0.416	0.654	0.197
Area 3—Subarea 4	0.156	0.203	0.279	0.309	0.018	0.028
Area 4—Subarea 2	0.0595	***0.126***	0.0925	***0.160***	0.0009	***0.0019***

In bold, we highlighted the smallest value among the different result in the column

3.4 Impact of outliers on the results

In order to assess the impact of outliers on the results, we applied an outlier removal method (less than 10% and more than 90%) and the outlier points were removed. We ran the model 10–20 times and considered the run with the lowest error, with the data after removing the outliers and we did not notice significant changes in the results. Indeed, in some cases, the results without the outliers were better (based on the MSE), while in others, the results were worse. Three examples are shown in Tables 13 through 15.

As can be seen, in Table 13, keeping the outlier data points resulted in higher MSE values and therefore worse predictions. When Area 4 was considered at both the complete area level and one of its subareas, i.e., Tables 14 and 15, the MSE values were lower when the outliers were kept as part of the dataset.

3.5 Discussion and practical cases

The experimental analysis demonstrates that the model's accuracy improves considerably when the training data percentage is increased. Specifically, increasing the training data from 70 to 90% led to a reduction in MSE across multiple areas, suggesting that the additional data improved the model's ability to learn the underlying patterns of ice thickness changes more

Table 13 Results with and without outliers for Area 3—Subarea 3

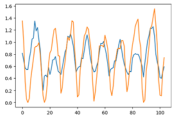
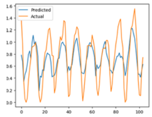
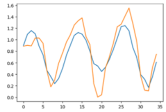
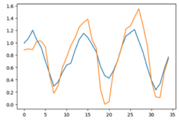
Data split	70–30%	70–30%	90–10%	90–10%
With outliers	Yes (1 point, < 1%)	No	Yes (1 point, < 1%)	No
MSE	0.089	0.077	0.0504	0.0501
Result				

Table 14 Results with and without outliers for Area 4

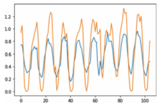
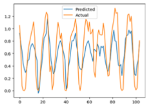
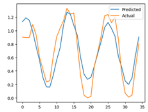
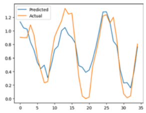
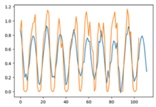
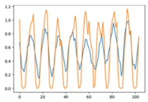
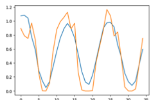
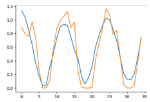
Data split	70–30%	70–30%	90–10%	90–10%
With outliers	Yes (1 point, < 1%)	No	Yes (1 point, < 1%)	No
MSE	0.068	0.0767	0.0379	0.043
Result				

Table 15 Results with and without outliers for Area 4—Subarea 2

Data split	70–30%	70–30%	90–10%	90–10%
With outliers	Yes (3 points, < 1%)	No	Yes (3 points, < 1%)	No
MSE	0.047	0.0598	0.028	0.0309
Result				

effectively. For instance, in Area 2, the MSE dropped from 0.118 (70% training) to 0.087 (90% training), while in Area 4, a similar reduction was noted from 0.068 to 0.0448. This indicates that the increased training data allowed the model to better generalize and make more accurate predictions. However, this improvement should be taken with care since this may be related to overfitting.

The model's performance was further evaluated based on seasonal inputs and varying prediction horizons. Notably, the model performed well for short-term predictions, achieving low MSE values when predicting up to 30–60 days in advance, with values of 0.0468 and 0.051, respectively. However, its performance deteriorated for longer prediction periods of 120 days, resulting in an MSE of 0.147. This indicates that the model is effective for short-term predictions but less reliable for mid- to long-term projections due to a reduced ability to capture longer-term variability.

In comparison to other studies, the results developed in this paper exhibit competitive accuracy. The use of LSTM neural networks for ice thickness prediction can be compared with the work of Wang et al. (2017), where Convolutional Neural Network (CNN) models were used to analyze Synthetic Aperture Radar (SAR) images for sea ice concentration. The results from that study also demonstrated improved accuracy by employing machine learning models, highlighting the effectiveness of deep learning for such complex prediction tasks. The error estimates by Wang et al. (2017) had Root Mean Squared Error (RMSE) values ranging between 0.1435 and 0.2214.

Moreover, other studies using ANN ensembles for ice thickness predictions on lakes (e.g., Zaier et al., 2010) have shown that combining outputs from separately trained models can lead to higher accuracy and stability in predictions. Zaier et al. (2010) have achieved RMSE values between 1.41 and 8.42 (cm) using different ANN ensemble configurations. This aligns with the findings of this paper using different data configurations, such as reducing the training period to more recent data, which helped in improving prediction performance due to reduced data variability.

The model's performance was also analyzed under different conditions, such as seasonal changes and the use of extreme data values. The experiments indicated that focusing on recent months instead of the entire dataset led to a significant reduction in prediction errors. For instance, reducing the input data to the recent 100 months and using 90% of this data for training reduced the MSE from 0.089 to 0.0276 for Area 2. This implies that seasonal variations, as well as changes in data trends over long periods, can negatively impact the model's performance unless appropriately managed.

The model struggled with areas that had high variability in ice thickness, as observed in Area 2, where the high standard deviation of ice thickness values resulted in reduced accuracy. Conversely, for areas like Area 4, where the variability was lower, the model achieved better results with an MSE of 0.0448. This highlights that the degree of variability in ice thickness plays a critical role in determining the accuracy of the predictions, suggesting that the model benefits from stable, less fluctuating datasets.

In scenarios involving extreme weather conditions, the decision to retain outlier values in the dataset was justified by the potential impact of extreme ice thickness values on the predictions. Outliers, such as very thick ice during specific seasons, can have a significant impact on navigation and safety in the Arctic, making their inclusion essential for realistic forecasting.

Moreover, an analysis of the impact of these outliers on the results showed that no significant difference can be noticed in the accuracy. This may maybe related to the fact that these data points correspond to recent changes that have been seen in Arctic weather due to climate change.

The findings suggest that increasing the training data percentage, focusing on recent data, and accounting for variability across areas are effective strategies for improving model accuracy. Compared to other studies, this model shows promise for short-term predictions but struggles with extended horizons and high variability. By retaining outliers and managing variability, the model achieves greater prediction robustness, which is crucial in real-world applications involving extreme environmental conditions. The results emphasize the need for tailored approaches depending on the specific characteristics of the data, such as seasonality and variability.

As stressed in the introduction, ice thickness has a direct impact on the navigability of vessels within the Arctic, and thus the capacity to predict the ice conditions is paramount. As an example, Fedi et al (2020) emphasized the impact that ice conditions had on three accidents with casualties in the Russian Arctic and highlighted the difficulty, yet the importance, of predicting ice conditions to provide safe navigation conditions. Among the three examples used by Fedi et al (2020), the case of the Sinegorsk could have been avoided if ice thickness had been predicted properly. This case demonstrates that a more precise and less extensive analysis of the ice thickness would have made it possible to avoid this accident. From an economic point of view, the anticipation of ice conditions may well increase the reliability of Arctic navigation for liner shipping. As explained by Cariou et al (2019), ice thickness has a negative impact on transit time, cost, and GHG emissions. Being able to forecast ice conditions may enable shipping companies to update information to shippers on the potential delay of containers or commodities.

4 Conclusion

Due to recent rapid changes in Arctic climate, a continuous decline of multiyear ice and ice thickness is making the Arctic Ocean a valid option for maritime navigation and shipping. Having better knowledge of sea ice thickness will make it easier to choose the routes for vessels and make navigation safer. This study examined the possibility of developing an ANN model to predict the sea ice thickness in the Arctic area. To achieve that, we chose four areas of interest and downloaded data from the Copernicus database for the sea ice parameters across 1991–2019. Data was preprocessed and transformed from NetCDF type to a more structured tabular form that could be used with the ANN. LSTM, a special type of RNN, was used to predict ice thickness values, as this type performs very well in such problems.

Our results emphasize that if the LSTM is a relevant tool to predict ice thickness, some parameters must be considered. First, the length of the prediction period with good accuracy depends on the type of data used (daily or monthly). Second, increasing the training percentage provides more accurate results. Third, when the variability of the amplitude of the ice thickness data, and when the variabilities of the maximum ice thickness value and the minimum ice thickness value are low, the LSTM model will provide more accurate results.

This study has some limitations. First, the obtained results, although they may be generalizable by collecting more data from other areas and testing the model, are limited to the areas that were considered. Second, the obtained results used only one ANN methodology, namely the LSTM, while other techniques may be used. Third, the model accuracy can be further improved by using more data or by relying on other techniques for hyperparameter optimization such as genetic algorithms. Fourth, the model was trained and used to predict only ice thickness, while other parameters may also be used such as ice concentration, wind speed, air temperature, etc., and then used to define the RIO of vessels based on POLARIS methodology. Predicting more parameters may allow a range of other applications such as fisheries management, oil and gas exploitation, infrastructure engineering projects, and search and rescue (SAR) operations. Finally, combining LSTM with other models for multimodal prediction, in addition to the previous points can also be seen as avenues for future research.

Declarations

Conflict of interest The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements) in the subject matter or materials discussed in this manuscript.

References

Abhinav, R., Pindoriya, N. M., Wu, J., & Long, C. (2017). Short-term wind power forecasting using wavelet-based neural network. *Energy Procedia, 142*, 455–460.

Abhishek, K., Singh, M. P., Ghosh, S., & Anand, A. (2012). Weather forecasting model using artificial neural network. *Procedia Technology, 4*, 311–318.

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon, 4*(11), e00938.

Aichouri, I., Hani, A., Bougherira, N., Djabri, L., Chaffai, H., & Lallahem, S. (2015). River flow model using artificial neural networks. *Energy Procedia, 74*, 1007–1014.

Aksenov, Y., Popova, E. E., Yool, A., Nurser, A. G., Williams, T. D., Bertino, L., & Bergh, J. (2017). On the future navigability of Arctic sea routes: High-resolution projections of the Arctic. *Marine Policy, 75*, 300–317.

Ali, S., Huang, Y., & Wang, J., et al. (2023). AI for sea ice forecasting. In Z. Sun (Ed.), *Artificial Intelligence in Earth Science* (pp. 41–58). Elsevier.

Bhoopathi, S., Kumar, N., & Pal, M. (2024). Evaluating the performances of SVR and XGBoost for short range forecasting of heatwaves across different temperature zones of India. *Applied Computing and Geosciences, 24*, 100204.

Borch, O.J., Westvik, M.H., Ehlers, S. & Berg, T.E. (2012). Sustainable arctic field and maritime operation. In *OTC Arctic Technology Conference,* Houston, Texas, USA, December 2012, Paper Number: OTC-23752-MS.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. (2015). *Time Series Analysis: Forecasting and Control*. Wiley.

Browne, T., Tran, T. T., Veitch, B., Smith, D., Khan, F., & Taylor, R. (2022). A method for evaluating operational implications of regulatory constraints on Arctic shipping. *Marine Policy, 135*, 104839.

Brownlee, J. (2017). Gentle Introduction to the Adam Optimization Algorithm for Deep Learning, Machine Learning Mastery, https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/, Accessed on 21 September 2020.

Bushaev, V. (2018). Adam — latest trends in deep learning optimization, Towards Data Science, https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c. Accessed on 22 September 2020.

Cabaneros, S. M. S., Calautit, J. K. S., & Hughes, B. R. (2017). Hybrid artificial neural network models for effective prediction and mitigation of urban roadside NO_2 pollution. *Energy Procedia, 142*, 3524–3530.

Cariou, P., Cheaitou, A., Faury, O., & Hamdan, S. (2019). The feasibility of Arctic container shipping: The economic and environmental impacts of ice thickness. *Maritime Economics & Logistics, 23*, 615–631.

Cheaitou, A., Faury, O., Cariou, P., Hamdan, S., & Fabbri, G. (2023). Economic and environmental impacts of Arctic shipping: A probabilistic approach. *Transportation Research Part d: Transport and Environment, 89*, 102606.

Cheridito, P., Jentzen, A., & Rossmannek, F. (2021). Non-convergence of stochastic gradient descent in the training of deep neural networks. *Journal of Complexity, 64*, 101540.

Copernicus. (2022). TOPAZ4 data: Arctic_Reanalysis_Phy_002_003, http://marine.copernicus.eu/services-portfolio/access-to-products/. Accessed 1 September 2020.

Copernicus. (2024). Arctic Ocean Physics Reanalysis data: ARCTIC_MULTIYEAR_PHY_002_003, https://data.marine.copernicus.eu/product/ARCTIC_MULTIYEAR_PHY_002_003/description Accessed 08 November 2024.

Copernicus Marine Services. (2024). Product User Manual For Arctic Ocean Physical and Bio Analysis and Forecasting Products. https://marine.copernicus.eu/sites/default/files/product_improvement_migrated_files/CMEMS-ARC-PUM-002-ALL.pdf. Accessed 09 November 2024.

de CarvalhoPaulino, Â., Guimarães, L. N. F., & Shiguemori, E. H. (2019). Hybrid Adaptive Computational Intelligence-based multisensor data fusion applied to real-time UAV autonomous navigation. *Inteligencia Artificial, 22*(63), 162–195.

Di, K., Chen, W., Zhang, X., Shi, Q., Cai, Q., Li, D., Liu, C., & Di, Z. (2023). Regional unevenness and synergy of carbon emission reduction in China's green low-carbon circular economy. *Journal of Cleaner Production, 420*, 138436.

Di, K., Chen, W., Shi, Q., Cai, Q., & Liu, S. (2024a). Analyzing the impact of coupled domestic demand dynamics of green and low-carbon consumption in the market based on SEM-ANN. *Journal of Retailing and Consumer Services, 79*, 103856.

Di, K., Chen, W., Shi, Q., Cai, Q., & Zhang, B. (2024b). Digital empowerment and win-win co-operation for green and low-carbon industrial development: Analysis of regional differences based on GMM-ANN intelligence models. *Journal of Cleaner Production, 445*, 141332.

Diebold, F. X., Göbel, M., Goulet Coulombe, P., Rudebusch, G. D., & Zhang, B. (2021). Optimal combination of Arctic sea ice extent measures: A dynamic factor modeling approach. *International Journal of Forecasting, 37*, 1509–1519.

Dong, Z., Shi, L., Lin, M., & Zeng, T. (2022). A suitable retrieval algorithm of Arctic snow depths with AMSR-2 and its application to sea ice thicknesses of cryosat-2 data. *Remote Sensing, 14*, 1041.

Duan, K., Huang, F., Zhang, S., Shu, Y., Dong, S., & Liu, M. (2024). Prediction of ship following behavior in ice-covered waters in the Northern Sea Route based on hybrid theory and data-driven approach. *Ocean Engineering, 296*, 116939.

Faury, O., Cheaitou, A., & Givry, P. (2020). Best maritime transportation option for the Arctic crude oil: A profit decision model. *Transportation Research Part e: Logistics and Transportation Review, 136*, 101865.

Fedi, L., Faury, O., & Gritsenko, D. (2018a). The impact of the Polar Code on risk mitigation in Arctic waters: A "toolbox" for underwriters? *Maritime Policy & Management, 45*(4), 478–494.

Fedi, L., Etienne, L., Faury, O., Rigot-Muller, P., Stephenson, S., & Cheaitou, A. (2018b). Arctic navigation: Stakes, benefits and limits of the POLARIS system. *The Journal of Ocean Technology, 13*(4), 54–67.

Fedi, L., Faury, O., & Etienne, L. (2020). Mapping and analysis of maritime accidents in the Russian Arctic through the lens of the Polar Code and POLARIS system. *Marine Policy, 118*, 103984.

Fu, S., Zhang, D., Montewka, J., Yan, X., & Zio, E. (2016). Towards a probabilistic model for predicting ship besetting in ice in Arctic waters. *Reliability Engineering & System Safety, 155*, 124–136.

Fu, S., Zhang, D., Montewka, J., Zio, E., & Yan, X. (2018). A quantitative approach for risk assessment of a ship stuck in ice in Arctic waters. *Safety Science, 107*, 145–154.

Gholami, V. C. K. W., Chau, K. W., Fadaee, F., Torkaman, J., & Ghaffari, A. (2015). Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. *Journal of Hydrology, 529*, 1060–1069.

Guo, H., Cui, Q., Shi, L., Parol, J., AlSanad, S., & Wu, H. (2024). Ice coating prediction based on two-stage adaptive weighted ensemble learning. *Processes, 12*(9), 1854.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.

Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in python. *Software Metapapers, 5*, 10.

Humpert, M. (2014). New analysis on the Northern Sea Route, https://www.highnorthnews.com/en/new-analysis-northern-sea-route. Accessed on 08 November 2024.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). Otexts.

IMO. (2008). Resolution MSC.255(84) adopted on 16 May 2008 adoption of the Code of the international standards and recommended practices for a safety investigation into a marine casualty or marine incident (Casualty Investigation Code).

IMO. (2014). Resolution MSC 385 (94) of November 21, 2014 and Resolution MEPC 264 (68) of May 15, 2015.International Code for Ships Operating in Polar Waters (Polar Code), 2014.

IMO. (2016). MSC. 1/Circ 1519, 6 June 2016, Guidance on Methodologies forAssessing Operational Capabilities and Limitations in Ice. https://docs.imo.org/Final Documents/English/MSC.1-CIRC.1519(E).docx

Karevan, Z., & Suykens, J. K. (2020). Transductive LSTM for time-series prediction: An application to weather forecasting. *Neural Networks, 125*, 1–9.

Khrapov, V. & Yushchenko, V. (2019). Logistics in the Arctic Region through the example of cargo transportation in Murmansk Oblast by railway transport. In: IOP Conference Series (Eds.) Earth and Environmental Science. IOP Publishing, Vol. 302, No. 1, 012107.

Kırbaş, İ, Sözen, A., Tuncer, A. D., & Kazancıoğlu, F. Ş. (2020). Comperative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons & Fractals, 138*, 110015.

Koyama, T., Nakanowatari, T., & Inoue, J. (2021). Information retrieval for Northern Sea Route (NSR) navigation: A statistical approach using the AIS and TOPAZ4 data. *Polar Science, 27*, 100626.

Kozmenko, S., Teslya, A. & Fedoseev, S. (2018). Maritime economics of the Arctic: Legal regulation of environmental monitoring. In: IOP conference series (Eds.) Earth and Environmental Science. IOP Publishing, Vol. 180, No. 1, 012009.

Lee, S. W., & Song, J. M. (2014). Economic possibilities of shipping though Northern Sea route. *The Asian Journal of Shipping and Logistics, 30*(3), 415–430.

Li, P., Leng, H., Ma, B. & He, M. (2021).Monthly Arctic sea ice concentration prediction based on the ConvLSTM Model. In: CISAI (Eds.). Kunming, China, September 17–19, 2021.

Li, W., Hsu, C. Y., & Tedesco, M. (2024). Advancing Arctic sea ice remote sensing with AI and deep learning: Now and future. *Egusphere, 2024*, 1–36.

Lindstad, E., Bright, R. M., & Stromman, A. H. (2016). Economic savings linked to future Arctic shipping trade are at odds with climate change mitigation. *Transport Policy, 45*(22), 24–30.

Liu, J., Chen, Z., Hu, Y., Zhang, Y., Ding, Y., Cheng, X., Yang, Q., Nerger, L., Spreen, G., Horton, R., & Inoue, J. (2019). Towards reliable Arctic sea ice prediction using multivariate data assimilation. *Science Bulletin, 64*(1), 63–72.

Liu, Q., Zhang, R., Wang, Y., Yan, H., & Hong, M. (2021). Daily prediction of the Arctic sea ice concentration using reanalysis data based on a convolutional LSTM network. *Journal of Marine Science and Engineering, 9*, 330.

Liu, L., Davedu, S., Fujisaki-Manome, A., Hu, H., Jablonowski, C., & Chu, P. Y. (2022). Machine learning model-based ice cover forecasting for a vital waterway in large lakes. *Journal of Marine Science and Engineering, 10*, 1022.

Liu, M., Wang, Y., Xing, Z., Wang, X., & Fu, Q. (2023). Study on forecasting break-up date of river ice in Heilongjiang province based on LSTM and CEEMDAN. *Water, 15*, 496.

Madan, T., Sagar, S., Tran, T. A., Virmani, D., & Rastogi, D. (2025). Air quality prediction using ensemble classifiers and single decision tree. *Journal of Advanced Research in Applied Sciences and Engineering Technology, 52*(1), 56–67.

Marchenko, N. (2014). Northern Sea Route: Modern State and Challenges. Proceedings of ASME $2^{01}4$ 33rd International Conference on Ocean, Offshore and Arctic Engineering (pp. V010T07A022-V010T07A022). American Society of Mechanical Engineers.

Marín-Blázquez, J.G. & Schulenburg, S. (2006). Multi-step environment learning classifier systems applied to hyper-heuristics. In: Proceedings of the 8th annual conference on Genetic and evolutionary computation (Eds.), pp. 1521–1528.

McLeod, S. (2019). What does a box plot tell you?, https://www.simplypsychology.org/boxplots.html. Accessed on 20 November 2020.

Meier, W. N., & Stroeve, J. (2022). An updated assessment of the changing Arctic sea ice cover. *Oceanography, 35*(3/4), 10–19.

Milaković, A. S., Ehlers, S., Westvik, M. H., & Schütz, P. (2014). Offshore upstream logistics for operations in Arctic environment. *Sun above the Horizon: Meteoric Rise of the Solar Industry, 5*, 163.

Murphy, A. Batch Size (Machine Learning), https://radiopaedia.org/articles/batch-size-machine-learning. Accessed on 10 November 2017.

Niedbała, G. (2019). Simple model based on artificial neural network for early prediction and simulation winter rapeseed yield. *Journal of Integrative Agriculture, 18*(1), 54–61.

NSR Information Office, Center for High North Logistics, https://arctic-lio.com/. Accessed on 20 December 2020.

Ortega, L. C., Otero, L. D., Solomon, M., Otero, C. E., & Fabregas, A. (2023). Deep learning models for visibility forecasting using climatological data. *International Journal of Forecasting, 39*, 992–1004.

Phutthaphaiboon, T., Siripongwutikorn, P. & Pusawiro, P. (2023). Forecasting Arctic Sea Ice Concentration using Long Short-term Memory Networks. In: ICMLT (Eds.). Stockholm, Sweden, March 10–12, 2023.

Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy, 148*, 461–468.

Ressel, R., Frost, A., & Lehner, S. (2015). A neural network-based classification for sea ice types on X-band SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(7), 3672–3680.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., & Haas, C. (2017). A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data. *The Cryosphere, 11*(4), 1607–1623.

Rigot-Müller, P., Cheaitou, A., Etienne, L., Faury, O., & Fedi, L. (2022). The role of Polarseaworthiness in shipping planning for infrastructure projects in the Arctic: The case of Yamal LNG plant. *Transportation Research Part a: Policy and Practice, 155*, 330–353.

SadefoKamdem, J., MianoMukami, P., & Njong, J. B. (2023). Time-frequency analysis and machine learning models for carbon market forecasting. *Annals of Operations Research*. https://doi.org/10.1007/s10479-023-05443-x

Shahzad, U., Sengupta, T., Rao, A., & Cui, L. (2023). Forecasting carbon emissions future prices using the machine learning methods. *Annals of Operations Research*. https://doi.org/10.1007/s10479-023-05188-7

Sun, T., Di, K., & Shi, Q. (2024). Digital economy and carbon emission: The coupling effects of the economy in Qinghai region of China. *Heliyon, 10*(4), e26451.

Theocharis, D., Pettit, S., Sanchez Rodrigues, V., & Haider, J. (2018). Arctic shipping: A systematic literature review of comparative studies. *Journal of Transport Geography, 69*, 112–128.

Theocharis, D., Rodrigues, V. S., Pettit, S., & Haider, J. (2019). Feasibility of the Northern Sea Route: The role of distance, fuel prices, ice breaking fees and ship size for the product tanker market. *Transportation Research Part e: Logistics and Transportation Review, 129*, 111–135.

Umamaheswari, P., & Ramaswamy, V. (2024). An integrated framework for rainfall prediction and analysis using a Stacked Heterogeneous Ensemble Model (SHEM). *Expert Systems with Applications, 256*, 124831.

Venkatachalam, K., Trojovský, P., Pamucar, D., Bacanin, N., & Simic, V. (2023). DWFH: An improved data-driven deep weather forecasting hybrid model using Transductive Long Short Term Memory (T-LSTM). *Expert Systems with Applications, 213*, 119270.

Wang, L., Scott, K. A., & Clausi, D. A. (2017). Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network. *Remote Sensing, 9*(5), 408.

Wang, K., Tang, X. Y., & Zhao, S. (2024a). Robust multi-step wind speed forecasting based on a graph-based data reconstruction deep learning method. *Expert Systems with Applications, 238*, 121886.

Wang, Q., Liu, H., Li, Y., Li, W., Sun, D., Zhao, H., Tie, C., Gu, J., & Zhao, Q. (2024b). Predicting plateau atmospheric ozone concentrations by a machine learning approach: A case study of a typical city on the southwestern plateau of China. *Environmental Pollution, 363*, 125071.

Wang, J., Qiao, L., Zhu, G., Di, K., & Zhang, X. (2025). Research on the driving factors and impact mechanisms of green new quality productive forces in high-tech retail enterprises under China's dual carbon goals. *Journal of Retailing and Consumer Services, 82*, 104092.

Wei, J., Hang, R., & Luo, J. J. (2022). Prediction of Pan-Arctic sea ice using attention-based LSTM neural networks. *Frontiers in Marine Science, 9*, 860403.

Wu, H., Wang, Y., Zhang, R., Yan, H., & Hong, M. (2024). Bias correction of Arctic sea ice thickness products based on factor selection and machine learning methods. *Applied Ocean Research, 149*, 104069.

Xue, H., Cai, M., Liu, B., Di, K., & Hu, J. (2024). Sustainable development through digital innovation: Unveiling the impact of big data comprehensive experimental zones on energy utilization efficiency. *Sustainable Development*. https://doi.org/10.1002/sd.3112

Yang, Y., Gao, P., Sun, Z., Wang, H., Lu, M., Liu, Y., & Hu, J. (2023). Multistep ahead prediction of temperature and humidity in solar greenhouse based on FAM-LSTM model. *Computers and Electronics in Agriculture, 213*, 108261.

Yang, X., Lin, Z. Y., Zhang, W. J., Xu, S., Zhang, M. Y., Wu, Z. D., & Han, B. (2024). Review of risk assessment for navigational safety and supported decisions in arctic waters. *Ocean & Coastal Management, 247*, 106931.

Yibre, A. M., & Koçer, B. (2021). Semen quality predictive model using feed forwarded neural network trained by learning-based artificial algae algorithm. *Engineering Science and Technology, an International Journal, 24*, 310–318.

Zaatar, T., Cheaitou, A. & Faury, O. (2021). Arctic ice thickness prediction using artificial neural networks. In: 2021 14th International Conference on Developments in eSystems Engineering (Eds.). IEEE: Sharjah, UAE, December 07–10 2021.

Zaier, I., Shu, C., Ouarda, T. B. M. J., Seidou, O., & Chebana, F. (2010). Estimation of ice thickness on lakes using artificial neural network ensembles. *Journal of Hydrology, 383*(3–4), 330–340.

Zeng, N., Sun, J., & Yu, L. (2014). Decisions of shipping supply chain with Arctic route option and emission allowance. *Transportation Research Part d: Transport and Environment, 136*, 104420.

Zhang, C., Zhang, D., Zhang, M., Lang, X., & Mao, W. (2020). An integrated risk assessment model for safe Arctic navigation. *Transportation Research Part a: Policy and Practice, 142*, 101–114.

Zheng, Q., Li, W., Shao, Q., Han, G., & Wang, X. (2022). A mid- and long-term Arctic sea ice concentration prediction model based on deep learning technology. *Remote Sensing, 14*, 2889.

Zhou, M., Lu, W., Ma, Q., Wang, H., He, B., Liang, D., & Dong, R. (2023). Study on the snowmelt flood model by machine learning method in Xinjiang. *Water, 15*, 3620.