

Article

Analysis of Dynamic Changes in Sea Ice Concentration in Northeast Passage during Navigation Period

Yawen He ¹, Yanhua Liu ^{1,*}, Duxian Feng ², Yongheng Li ¹, Feng Jin ¹ and Jinxiu Deng ¹¹ College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China; heyw@upc.edu.cn (Y.H.)² CPECC Xinjiang Petroleum Engineering Co., Ltd., Karamay 834000, China

* Correspondence: z22160029@s.upc.edu.cn

Abstract: With global warming and the gradual melting of Arctic sea ice, the navigation duration of the Northeast Passage (NEP) is gradually increasing. The dynamic changes in sea ice concentration (SIC) during navigation time are a critical factor affecting the navigation of the passage. This study uses multiple linear regression and random forest to analyze the navigation windows of the NEP from 1979 to 2022 and examines the critical factors affecting the dynamic changes in the SIC. The results suggest that there are 25 years of navigable windows from 1979 to 2022. The average start date of navigable windows is approximately between late July and early August, while the end date is approximately early and mid-October, with considerable variation in the duration of navigable windows. The explanatory power of RF is significantly better than MLR, while LMG is better at identifying extreme events, and RF is more suitable for assessing the combined effects of all variables on the sea ice concentration. This study also found that the 2 m temperature is the main influencing factor, and the sea ice movement, sea level pressure and 10 m wind speed also play a role in a specific period. By integrating traditional statistical methods with machine learning techniques, this study reveals the dynamic changes of the SIC during the navigation period of the NEP and identifies its driving factors. This provides a scientific reference for the development and utilization of the Arctic Passage.



Citation: He, Y.; Liu, Y.; Feng, D.; Li, Y.; Jin, F.; Deng, J. Analysis of Dynamic Changes in Sea Ice Concentration in Northeast Passage during Navigation Period. *J. Mar. Sci. Eng.* **2024**, *12*, 1723. <https://doi.org/10.3390/jmse12101723>

Academic Editor: Anatoly Gusev

Received: 4 August 2024

Revised: 13 September 2024

Accepted: 23 September 2024

Published: 1 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Northeast Passage; sea ice; Arctic; Northern Sea route; LMG; random forest; importance measure

1. Introduction

As a critical channel within the Arctic Passage, the NEP connects the North Pacific and the North Atlantic. Its starting point is Russia, crossing the Arctic Ocean eastwards through the Barents Sea, the Kara Sea, the Laptev Sea, the East Siberian Sea and the Chukchi Sea, finally reaching the Bering Strait [1]. The benefits of the NEP include significantly reduced navigation time, cost savings and the avoidance of high-risk areas associated with traditional waterways [2].

However, due to the greenhouse effect and ocean heating caused by global climate change, the coverage and thickness of Arctic sea ice are decreasing, which brings new opportunities for the opening of the Arctic Passage [3]. Due to the existence of Arctic sea ice, the Northeast Passage is not navigable throughout the year, but there is a certain navigation period [4]. Li et al. [5], based on the Arctic sea ice dataset obtained from 1988 to 2020, the spatiotemporal trends of Arctic SIC are analyzed. Vyacheslav C et al. [6] developed a new method to calculate the characteristics of transit navigation between Atlantic and Pacific regions along the NSR. Kibanova et al. [7] analyze the changes in the NPD on the NSR for the CMIP5 climate model ensemble under RCP anthropogenic scenarios using Bayesian statistics. Vladimir [8] points to the changes in energy fluxes in the polar ocean–ice–atmosphere system. The rising concentrations of atmospheric greenhouse gases are

viewed as a main driving factor of the sea ice decline trend [9]. With the preferential loss of multi-year ice in comparison with relatively thin first-year ice, the downward trend of the Arctic sea ice extent steepens from the end of the 20th century [10]. However, most of the current climate models do not correctly capture this significant decrease [11]. This indicates that our understanding of the controlling factors of sea ice melt may still be insufficient, especially in a warmer climate.

Under the background of studying the impact of climate change on the navigation of the NEP, it is very crucial to understand and quantify the relative importance of influencing factors [12]. This not only aids in predicting and assessing the changes in the navigable window but also provides a scientific basis for shipping planning and policy formulation. Thermodynamic factors such as a 2 m temperature [13] and 10 m wind speed [14], as well as dynamic factors like sea ice motion [15], significantly impact the navigation conditions of the channel. By quantifying the relative importance of these factors, researchers can more accurately identify which factors have a decisive impact on navigability under specific periods and conditions, thereby optimizing shipping decisions [16].

This study utilizes the variable importance measure (VIM) [17] in the regression model to examine the relative importance of various factors and their detailed relationship with the dynamic changes in the SIC. VIM helps better understand the process that may generate data and addresses questions related to the most relevant factors that influence the outcomes [18]. Multiple linear regression (MLR) analysis is one of the most commonly used methods to study the relationship between the dynamic changes in the SIC and related factors. However, most MLRs cannot directly evaluate the importance of variables. Many studies use correlation coefficients and regression coefficients to evaluate the importance of variables, but these methods can lead to unstable and misleading results when variables exhibit collinearity [19]. Various advanced statistical methods have been developed to measure variable importance, including variance decomposition, variable transformation and machine learning algorithms [20]. The method based on variance decomposition is most widely applied in MLR [21]. These methods decompose the coefficient of determination (R^2) into non-negative contributions attributable to each explanatory variable based on an average ranking [22]. Specifically, the parametric regression methods based on variance decomposition include Lindeman, Merenda and Gold (LMG) [23], dominance analysis [24] and proportional marginal variance decomposition [25]. The random forest (RF) algorithm, proposed by Breiman [26] in 2001, is a nonparametric regression technique that also has a specific method to derive variable importance.

In climate and environmental science research, MLR [27] and random forest (RF) [28] are commonly used analytical methods. MLR is a traditional statistical method that evaluates the importance of independent variables by establishing a linear relationship model between independent and dependent variables. It is computationally simple, highly interpretable and widely applied across various scientific fields. However, MLR assumes a linear relationship between variables, which may be limiting when dealing with complex climate data [29]. Random forest is a nonlinear method that makes predictions and assesses variable importance by constructing an ensemble of decision trees. RF can not only deal with nonlinear relationships but also capture complex interactions between variables [30]. In future research, combining the advantages of these two methods, more comprehensive and accurate analysis results are provided to provide more powerful support for the navigation prediction and planning of the NEP.

The primary objective of this study is to analyze the changes in the SIC and its influencing factors during the navigation period of the NEP by constructing MLR and RF. The specific research content includes collecting and organizing the SICs and related meteorological data from 1990 to 2022; the variance inflation factor (VIF) method was used to screen the independent variables, eliminate the influence of multicollinearity and ensure the stability and reliability of the model. MLR and RF were constructed to evaluate the relative importance of each meteorological variable to the SIC. By combining statistical methods and machine learning techniques, this study aims to reveal the dynamic changes

in the SIC and its driving factors during the navigation period of the NEP, providing scientific reference for the development and utilization of the Arctic Passage.

2. Research Data and Methods

2.1. Study Area

The Arctic region refers to the area north of latitude $66^{\circ}34'$ N (the Arctic Circle), encompassing most of the Arctic Ocean, as well as the coastal land and islands of Asia, Europe and North America. The Arctic climate is severe, with most areas covered by ice and snow. The Arctic Passage refers to a collection of sea lanes that cross the Arctic Ocean and connect the Pacific Ocean and the Atlantic Ocean. The navigation of the Arctic Passage is typically constrained by seasonal and ice conditions. The Arctic Passage can be divided into the NEP, the Northwest Passage and the Arctic Ocean Central Passage [31]. This paper mainly focuses on the NEP in the Arctic Passage and the associated maritime regions.

The Northeast Passage (NEP) is a maritime passage connecting the North Pacific and the North Atlantic. It starts from Russia and then heads east from Northern Europe and goes eastward through the Norwegian Sea, the Barents Sea, the Kara Sea, the Laptev Sea, the East Siberian Sea and the Chukchi Sea in the Arctic Ocean to the Bering Strait (Figure 1). As a critical channel within the Arctic Passage, the NEP has the advantages of shortening navigation times, saving navigation costs, and avoiding high-risk areas [32]. With the accelerated melting of Arctic sea ice, the importance and feasibility of navigation in the NEP are continuously increasing.

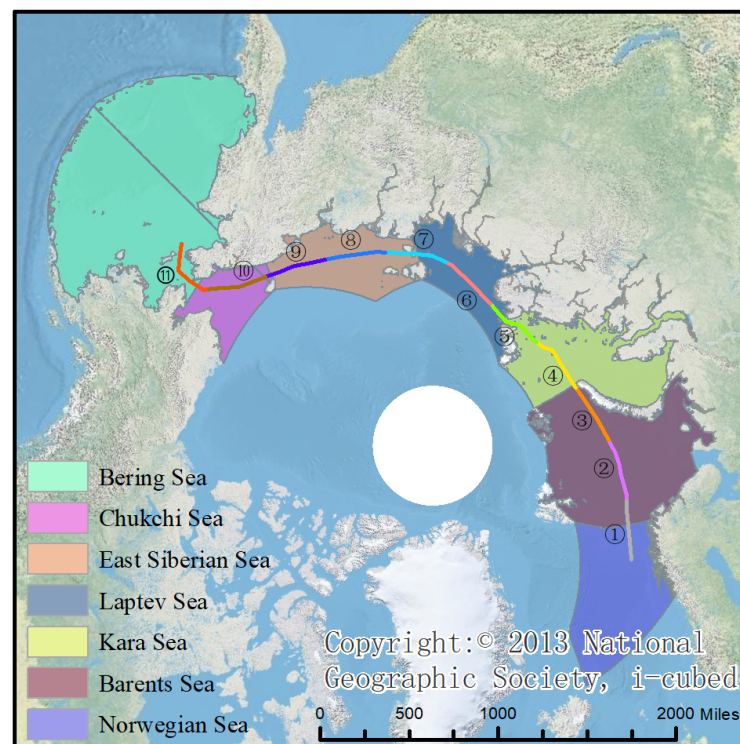


Figure 1. The schematic diagram of the study area and the key sea area (① to ⑪) indicates that the NEP is divided into 11 sections, which are distinguished by different colors. ① denotes the distance sailed on the first day of navigation, and so on).

2.2. Research Data

The data used in this study include the sea ice concentration (SIC) data, supplemented by the summer Arctic 2 m temperature, 10 m wind speed, sea level pressure and sea ice motion for the analysis.

The daily SICs released by the National Snow and Ice Data Center (NSIDC) were used as the basic dataset for this study. This dataset is derived from the SIC of Nimbus-7 SMMR

and DMSP SSM/I-SSMIS passive microwave data. In version 2 (<https://nsidc.org/data/nsidc-0051>, accessed on 15 February 2023), the dataset uses the brightness temperature data provided by microwave remote sensing sensors (SMMR, SSM/I, SSMIS), and uses the NASA Team algorithm for inversion generation. The spatial resolution of the dataset is 25 km, with a temporal resolution of daily observations covering the period from October 1978 to the present [33]. This study uses data from 1979 to 2022.

The 2 m temperature, 10 m wind speed and sea level pressure were obtained from the ERA5 reanalysis gridded dataset provided by the Climate Reanalyzer tool of the University of Maine (https://climatereanalyzer.org/research_tools/monthly_maps/, accessed on 1 April 2024). The temporal resolution is monthly, and the grid resolution is $0.25^\circ \times 0.25^\circ$. The dataset spans from 1979 to the present, and this study uses data from 1990 to 2022. The 10 m wind speed is the horizontal speed of the wind, or the movement of the air, measured at a height of ten meters above the Earth's surface, with units in meters per second (m/s). The 2 m temperature is the temperature at which the air, at the height of two meters above the Earth's surface, must be cooled to reach saturation, with units in Kelvin (K). The sea level pressure is the atmospheric pressure at the Earth's surface adjusted to the sea level height, with units in Pascals (Pa). Sea level pressure maps are used to identify the locations of low-pressure and high-pressure weather systems, commonly referred to as cyclones and anticyclones.

Sea ice motions were obtained from the Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors dataset, version 4, provided by the National Oceanic and Atmospheric Administration (NOAA) (<https://nsidc.org/data/nsidc-0116/versions/4>, accessed on 1 April 2024). The spatial resolution is 25×25 km, and the temporal resolution is 7 days. The dataset spans from 1978 to the present, and this study uses data from 1990 to 2022.

2.3. Research Method

2.3.1. Navigation Window Improvement Algorithm

The navigational period within the channel signifies the duration when the SIC does not jeopardize the safe passage of ships, thereby reflecting the SIC along the route of the channel. According to the description of SIC and sea ice navigation conditions by Shibata [34], when the SIC is 10–30%, there is only a small amount of broken ice on the sea surface, which is very sparse, and the traffic is very smooth at this time. When the SIC exceeds 40%, there is a significant amount of unconnected fragmented ice on the sea surface, making navigation more challenging. Therefore, in this study, grid cells with a SIC of less than 40% are considered navigable.

Based on the coordinates of the start and end points of the NEP and the cruising speeds of various vessels, this study assumes a vessel speed of 13 knots. Assuming continuous navigation along the NEP, this study stipulates that vessels can traverse the Arctic within these 11 days. Using a minimum transit period of 11 days as the condition, the navigation period is calculated by the improved navigation window algorithm, and the optimal channel route is determined according to the SIC in different regions. The core idea of this algorithm is to assess whether starting from a given day, there exists a path that ensures a vessel can traverse the Arctic within these 11 days.

2.3.2. Variance Inflation Factor (VIF)

Before constructing the regression model, it is necessary to assess multicollinearity among the variables and exclude those that are highly correlated. The variance inflation factor (VIF) is used to eliminate the multicollinearity. Initially, the full model containing all variables is constructed under the initial condition. According to the VIF value of each variable, when the VIF is greater than 10, the variable is removed, and this process is iteratively repeated until all variables have VIF values of less than 10. After eliminating multicollinearity, MLR is applied to measure the relative importance of each variable on the SICs in anomalous years.

The calculation formula of *VIF* is [35]:

$$VIF_j = \frac{1}{1 - R_j^2}, \quad (1)$$

where R_j represents the coefficient of determination obtained from the auxiliary regression of multiple explanatory variables, j represents the j th variable and VIF_j represents the variance inflation factor for calculating the j th independent variable.

The advantage of *VIF* is that it can assess multicollinearity and the multiple correlations among variables. The general criteria for *VIF* are as follows: when $0 < VIF < 5$, there is no multicollinearity; when $5 < VIF < 10$, there is weak multicollinearity; when $10 < VIF < 100$, there is moderate multicollinearity; when $VIF > 100$, there is severe multicollinearity [36].

2.3.3. Lindeman, Merenda and Gold (LMG) Index

The Lindeman, Merenda and Gold (LMG) index is calculated by considering the direct contribution of the variable and its marginal benefit when combined with all other variables [37]. It decomposes R^2 into non-negative contributions based on semi-biased coefficients [38] and comprehensively considers both the direct contribution of the variables and their marginal contributions after adjusting for other regression variables in the model.

When S variables are given, the R^2 of the MLR can be expressed as the ratio of the regression sum of squares of regression (*RSS*) to the total sum of squares (*TSS*) deviations:

$$R^2(S) = \frac{RSS}{TSS}, \quad (2)$$

where R^2 is a statistical index to measure the goodness-of-fit of the model; its value is between 0 and 1, which is used to represent the proportion of the variation explained by the model to the total variation. S represents the number of variables.

When a new variable M is added to the model, the additional increase in R^2 is defined as $seqR^2(M/S)$ and expressed as follows:

$$seqR^2(M/S) = R^2(M \cup S) - R^2(S), \quad (3)$$

where $seqR^2(M/S)$ here refers to how much the R^2 value of the whole model increases after adding the new variable M in the case of the existing variable S . The specific calculation formula is the R^2 value of the model after the new variable M is added (denoted by $R^2(M \cup S)$) minus the R^2 value when only the original variable S is added (denoted by $R^2(S)$). $seqR^2(M/S)$ represents the increment of the explanatory power of the variable M to the model. If this increment is large, the variable M contributes more to the improvement of the model.

Suppose that the order of variables $x_1, \dots, x_p$ in the model is $r = (r_1, \dots, r_p)$, $S_k(r)$ denotes the order of variables r before the variable x_k enters the model, then when x_k enters the model, the R^2 part assigned to x_k can be expressed as:

$$seqR^2(\{x_k\}/S_k(r)) = R^2(\{x_k\} \cup S_k(r)) - R^2(S_k(r)). \quad (4)$$

Then, for the explanatory variable x_k , the LMG index is expressed as follows:

$$LMG(x_k) = \frac{1}{p!} \sum_{r \text{ permutation}} seqR^2(\{x_k\}r), \quad (5)$$

where P denotes the total number of explanatory variables in the model. The symbol ‘!’ represents the factorial, that is, the factorial of p , representing the product of all integers from 1 to p . The $P!$ in the molecule represents the total number of permutations and combinations of all possible explanatory variables. Furthermore, “permutation” refers to the concept of permutation and combination, which means that considering the order of all

possible variable combinations, $seqR^2(\{x_k\} | r)$ represents the relative importance of variable x_k in the model in the case of a given variable order r .

The primary advantage of the LMG index is that it does not depend on the order in which each variable enters the model. Additionally, by utilizing both the direct contribution of a variable (x_k first enters the model) and its marginal contribution after adjusting for other variables, the LMG index provides a more accurate measure of each variable's overall contribution to the variance of the dependent variable [39].

2.3.4. Random Forest (RF) Variable Importance

Random forest (RF) is an ensemble learning method used for classification, regression and other tasks, which is based on multiple decision trees composed of several decision nodes. Compared with other machine learning methods, RF is more interpretable. For example, in comparison with support vector machines (SVMs), RF directly provides a feature importance evaluation, whereas SVM requires intricate mathematical derivations to obtain similar information. When handling high-dimensional data, SVM may suffer from the curse of dimensionality, while RF demonstrates more robust performance on high-dimensional datasets. Contrasted with K-nearest neighbors (K-NN), RF can effectively manage high-dimensional and large-scale datasets, where K-NN may perform sub-optimally. Additionally, RF typically trains faster than K-NN, as K-NN necessitates computing the distances between new data points and all training samples during prediction. Compared to naive Bayes, RF does not rely on the independence assumption of naive Bayes, thus making it more applicable to a wider range of data scenarios. Overall, RF possesses more intuitive, robust and easily interpretable advantages in feature importance evaluations compared to specific machine learning methods, rendering it a preferred tool in numerous data science tasks.

Within the framework of RF, the importance of the variable is measured by the increase in the prediction error calculated when the out-of-bag (OOB) data of the variable is permuted while all other data remain unchanged [40]. There are two main importance measures in RF: permutation-based and node impurity-based measures [41]. The permutation-based measure calculates the prediction accuracy of each tree on the OOB data (classification accuracy is expressed by the error rate, while regression accuracy is expressed by the mean squared error (MSE)) and then the prediction accuracy after permuting each variable. Node impurity is calculated by averaging over all variables, determining the total decrease in node impurity after the variable splits. In regression-focused RF algorithms, the most widely used importance measure is the permutation-based method [19].

Although RF possesses advantages over linear models and is more interpretable than other machine learning methods, its performance in ranking variable importance can be unstable. Strobl [42] found that importance scores can be biased when there is a correlation among variables. Genuer [43] conducted extensive theoretical and experimental simulation studies to explain the properties of the variable importance index in RF and proposed a stable method for determining the relative importance of explanatory variables. This method is based on the permutation-based measure. For each tree t in the forest, the OOB error is denoted as $erOOB_t$. When the variable X^j values in the OOB_t sample are randomly permuted, the resulting error is denoted as $er\widehat{OOB}_t^j$. Then, the variable importance of the variable X^j is expressed as:

$$VI(X^j) = \frac{1}{ntree} \sum_t \left(er\widehat{OOB}_t^j - erOOB_t \right) \quad (6)$$

where “ $ntree$ ” represents the number of trees, and ‘in the random forest’ refers to the total number of trees in the forest. j represents the j th variable.

In this algorithm, the steps of calculating the importance of the variables are mainly divided into the two following steps.

Step 1: The preliminary screening and ranking of variables. Initially, the raw variable importance is ranked by running the RF 50 times and taking the average importance scores. Subsequently, a threshold is set to eliminate the variables that are deemed to be of negligible importance.

Step 2: By establishing a series of nested RFs, each model includes the k variables listed in Step 1. When k ranges from 1 to m (assuming there are m total variables), the variables that minimize the OOB error are selected (after running the RF 25 times and taking the average). At this time, the m variables will be selected.

Ideally, the final model contains the variables that correspond to the minimum OOB error. In order to mitigate the impact of random instability in a RF, the algorithm selects the OOB error corresponding to the minimum standard deviation in 25 repeated experiments less than the OOB error corresponding to the minimum standard deviation in the repeated experiments.

The pseudocode is illustrated in Algorithm 1.

Algorithm 1: Random Forest

Input:

Training set $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$, features $F = \{f_1, f_2, \dots, f_m\}$ and number of trees in the forest B .

Output:

Ensemble of decision trees H

1 **function:** RANDOMFOREST(D, F)

2 Initialize H as an empty set

3 for $i = 1$ to B do

4 $S_i \leftarrow$ A bootstrap sample from D

5 $h \leftarrow$ RANDOMIZEDTREELEARN(S_i, F)

6 Add h to H

7 end for

8 return H

9 **function:** RANDOMIZEDTREELEARN(S, F)

10 At each node:

11 Randomly select a very small subset of features from F

12 Evaluate each feature in the subset using a splitting criterion (such as information gain or Gini impurity; this code uses Gini impurity.)

13 Choose the best feature based on the splitting criterion

14 Split the node based on the best feature

15 Recursively build the tree by repeating this process for each child node

16 Return the learned tree

This paper selected the parameters $ntree = 200$ (number of trees in RF model) and $mtry = 2$ (number of variables tried at each split in RF model) according to the guidance of the “random Forest” package in R. The experimental procedure is shown in Figure 2.

2.3.5. Relative Importance Analysis Process

In this study, the flow chart of the relative importance of the factors affecting the SIC using two models is shown in Figure 3.

All analyses in this paper use R3.6.2 software: r package ‘car 3.0–6’ is used to analyze the VIF and gradually eliminate variables with high collinearity with other explanatory variables; r package ‘relaimpo 2.2–3’ is used to calculate the LMG index; r package ‘VSURF’ is used to solve and deal with the variable importance of RF.

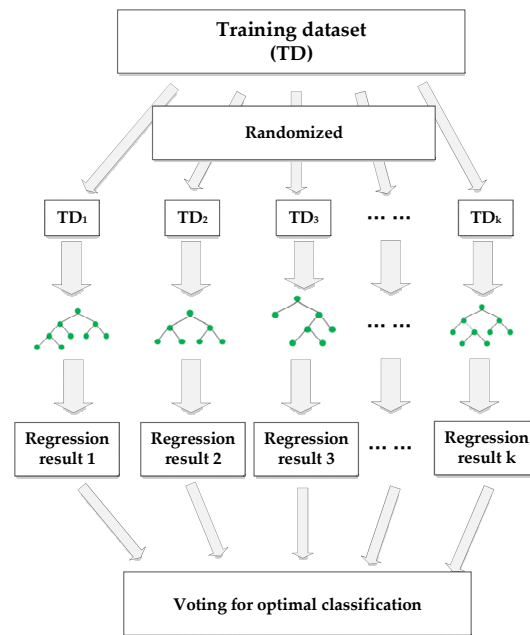


Figure 2. Structure of a random forest model.

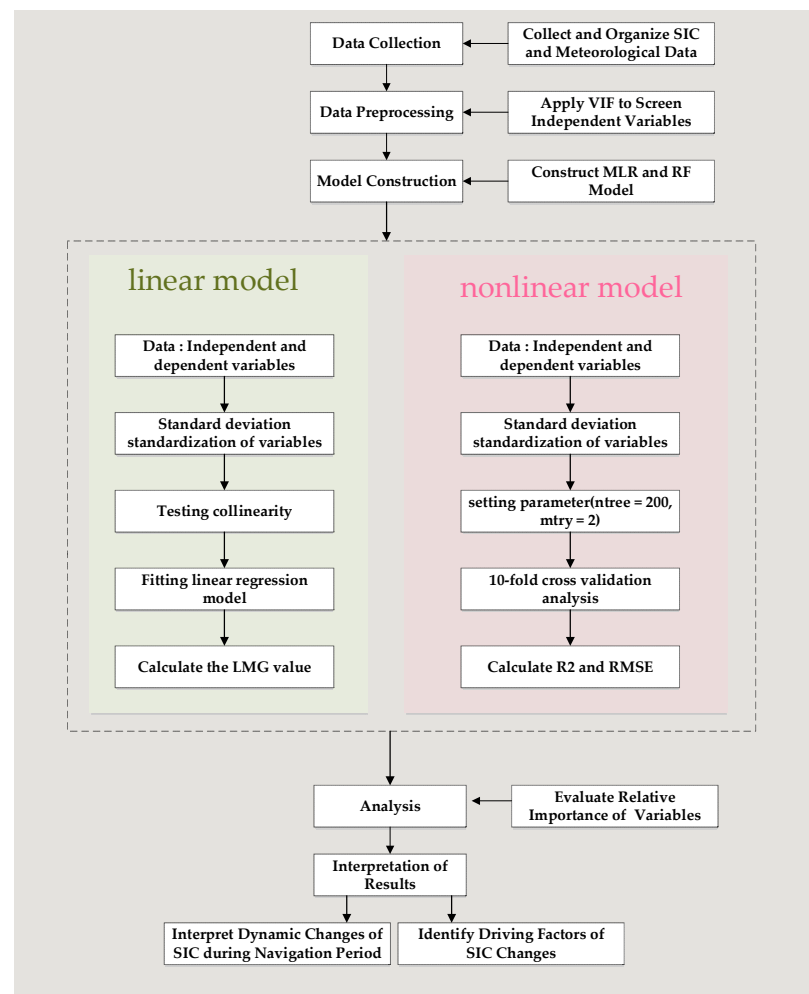


Figure 3. The flowchart of relative importance analysis.

3. Results

3.1. Navigation Window Calculation

This study analyzed the annual navigation start cycle, navigation end cycle and navigation cycle of the NEP from 1979 to 2022. The determination of navigability was based on the SIC in the relevant areas of the NEP. Specifically, a year was considered navigable if there existed a path through the NEP where the SIC was $\leq 40\%$, ensuring that a vessel could traverse the Arctic within 11 days (11 days was considered as one navigation cycle). From 1979 to 1989, the NEP was consistently covered with high concentrations of sea ice, rendering the route non-navigable during this period; hence, no navigation windows were recorded.

The basic situation of the navigation window period of the NEP from 1990 to 2022 is summarized in Table 1: there were 25 years with navigation windows between 1990 and 2022. The average starting time of navigation was approximately from late July to early August, while the average ending time was around early to mid-October. The average navigation period consisted of 60 days (excluding all non-navigable years). Within the 25 years that had navigation periods, there was a significant fluctuation in the duration, with no clear trend of a stable change.

Table 1. Northeast Passage navigation schedule 1979–2022.

Year	Navigation Start Cycle	Navigation End Cycle	Navigation Cycle	Year	Navigation Start Cycle	Navigation End Cycle	Navigation Cycle
1990	08–20	09–22	26	2007	08–05	10–04	60
1991	-	-	0	2008	08–11	10–01	52
1992	-	-	0	2009	08–13	10–11	56
1993	-	-	0	2010	08–04	10–03	61
1994	-	-	0	2011	07–17	10–13	89
1995	07–28	10–08	73	2012	07–25	10–19	82
1996	-	-	0	2013	08–09	10–01	53
1997	09–14	09–24	8	2014	07–29	10–05	69
1998	08–10	10–02	50	2015	07–17	10–10	86
1999	08–18	09–23	33	2016	08–11	10–17	68
2000	08–10	09–24	32	2017	07–25	10–07	72
2001	-	-	0	2018	07–29	10–19	83
2002	08–01	09–16	47	2019	07–07	10–12	96
2003	-	-	0	2020	07–07	10–28	112
2004	-	-	0	2021	07–30	10–02	64
2005	08–02	10–10	65	2022	08–12	10–08	57
2006	07–24	09–27	33				

Note: Navigation period: The total sum of all navigable periods within a year, represented by the start date of each individual navigable period. For example, in 1990, there were 26 navigable periods, with 08–22 representing the start date of the first period (08–22 to 09–01), indicating the first navigable period. ‘-’ means no value.

3.2. Results of Variance Inflation Factor (VIF)

The VIF was employed to exclude highly correlated variables, selecting all predictors with VIF values of less than 5. This process identified a total of four relevant variables. Additionally, the variables forecast albedo and sea ice age were excluded. The rationale for their exclusion stems from two key factors: firstly, their VIF values exceeded 5, and secondly, upon constructing a regression model involving six correlated factors alongside the SIC, the outcomes revealed that the influence of the 2 m temperature on the SIC accounts for more than 60%. Table 2 displays the VIF values of these four variables within the regression model.

Table 2. The VIF values of the four variables from 1990 to 2022.

Year	2 m Temperature	10 m Wind Speed	Sea Level Pressure	Sea Ice Motion
1990	1.25	1.57	1.32	1.57
1991	1.46	1.47	1.51	1.43
1992	1.38	1.12	1.42	1.05
1993	1.34	1.61	1.37	1.59
1994	1.62	1.44	1.57	1.50
1995	1.20	1.56	1.22	1.54
1996	1.71	1.29	1.73	1.15
1997	1.43	1.34	1.46	1.29
1998	1.40	1.34	1.34	1.41
1999	1.46	1.25	1.44	1.40
2000	1.42	1.30	1.44	1.31
2001	1.80	1.59	1.91	1.61
2002	1.34	1.23	1.37	1.21
2003	1.22	1.67	1.28	1.63
2004	1.33	2.80	1.38	2.76
2005	1.34	1.70	1.57	1.80
2006	1.33	1.36	1.25	1.42
2007	1.05	1.86	1.04	1.87
2008	1.06	1.61	1.03	1.65
2009	1.07	1.30	1.04	1.32
2010	1.04	1.50	1.01	1.51
2011	1.07	2.52	1.07	2.51
2012	1.19	1.03	1.07	1.17
2013	1.30	2.10	1.32	2.08
2014	1.21	1.15	1.11	1.22
2015	1.12	1.31	1.13	1.28
2016	1.15	1.48	1.18	1.44
2017	1.15	1.23	1.18	1.20
2018	1.18	2.02	1.10	2.07
2019	1.03	1.61	1.05	1.59
2020	1.19	1.14	1.05	1.21
2021	1.14	1.40	1.14	1.41
2022	1.73	1.39	2.84	2.18

3.3. Results of RF and LMG Effects of Factors on SIC in 2020

Table 3 shows the proportion of VIM values of the relevant elements from 1990 to 2022. In each year, the R^2 value is between 0.862 and 0.976, indicating that the variables considered have a good explanation for the dynamic changes in the SIC.

From 1990 to 2022, the influence of different related factors has also changed. For example, the contribution of the 2 m temperature varies between 33% and 78%, the contribution of the 10 m wind speed fluctuates between 5% and 26%, the contribution of sea ice drift varies between 20% and 57%, and the contribution of sea level pressure fluctuates between 1% and 31%. The contribution rate of a single related factor varies greatly, and the influence of the 2 m temperature and sea ice movement on the SIC accounts for more than half in some years. The linear model also shows that the effect of the 2 m temperature on the SIC ranks first.

The different calculation methods of the VIM and LMG lead to a difference in the data analysis results. The VIM focuses more on the overall impact of each independent variable, while the LMG focuses on the independent impact of each independent variable. Therefore, the changing trends of the VIM and LMG values may be different, which is reflected in some differences in the weight distribution of different elements.

Table 3. The proportion of VIM values of each variable from 1990 to 2022.

Year	R ²	2 m Temperature	10 m Wind Speed	Sea Level Pressure	Sea Ice Motion
1990	0.940	41%	20%	21%	18%
1991	0.938	43%	19%	21%	17%
1992	0.951	42%	14%	29%	15%
1993	0.960	42%	16%	28%	14%
1994	0.958	35%	24%	26%	15%
1995	0.926	37%	26%	26%	12%
1996	0.976	38%	20%	28%	14%
1997	0.940	41%	20%	27%	12%
1998	0.949	42%	20%	26%	13%
1999	0.925	40%	18%	27%	14%
2000	0.943	46%	17%	23%	14%
2001	0.961	41%	21%	25%	14%
2002	0.949	47%	21%	18%	14%
2003	0.937	41%	20%	23%	16%
2004	0.967	46%	25%	18%	11%
2005	0.888	40%	23%	18%	19%
2006	0.938	51%	15%	19%	15%
2007	0.862	46%	28%	10%	16%
2008	0.846	51%	29%	6%	14%
2009	0.718	46%	27%	4%	23%
2010	0.755	48%	27%	5%	19%
2011	0.883	42%	19%	21%	18%
2012	0.846	32%	17%	22%	29%
2013	0.934	44%	25%	16%	14%
2014	0.883	44%	24%	19%	13%
2015	0.905	43%	26%	17%	14%
2016	0.882	29%	23%	22%	27%
2017	0.879	44%	23%	20%	14%
2018	0.899	31%	36%	16%	16%
2019	0.834	39%	26%	16%	19%
2020	0.765	31%	19%	25%	25%
2021	0.907	32%	24%	16%	28%
2022	0.921	42%	24%	18%	16%

From 1990 to 2022, the proportion of VIM values of different related elements fluctuated, and the gap was small. For example, the proportion of the 2 m temperature varies between 39% and 51% in different years, the proportion of the 10 m wind speed fluctuates between 14% and 29%, the proportion of sea ice drift varies between 18% and 29% and the proportion of sea level pressure fluctuates between 11% and 29%. The effect of the 2 m temperature on the SIC ranks first among all related factors.

Table 4 shows the proportion of the LMG values of related factors from 1990 to 2022. These values represent the contribution of each variable in the overall model. The R² values ranged from 0.104 to 0.708, indicating that the explanatory power of the model varied greatly.

3.4. Effects of Factors on SIC in 2020

According to Table 1, 2020 is the longest year for navigation. Combined with the importance evaluation results of the VIM and LMG (Table 5 and Figure 4), the increase in the navigation period in 2020 is explained: both models show that the importance of the 2 m temperature increases in 2020, especially in the linear model LMG, indicating that the 2 m temperature has a significant impact on the navigation period in 2020, and the increase in high temperature also leads to a reduction in sea ice, thus extending the opening time of the navigation channel. The importance of sea ice movement also increased significantly in 2020, ranking second in the importance of sea ice movement in both models, indicating that a decrease in sea ice is also an important factor in the increase in navigation time.

Table 4. The proportion of LMG values of each variable from 1990 to 2022.

Year	R ²	2 m Temperature	10 m Wind Speed	Sea Level Pressure	Sea Ice Motion
1990	0.555	57%	11%	30%	2%
1991	0.569	57%	15%	26%	2%
1992	0.685	57%	7%	32%	4%
1993	0.693	48%	10%	41%	2%
1994	0.593	46%	26%	26%	2%
1995	0.463	52%	14%	29%	5%
1996	0.708	39%	16%	44%	1%
1997	0.579	46%	11%	42%	1%
1998	0.687	53%	14%	24%	9%
1999	0.663	49%	13%	31%	6
2000	0.582	52%	5%	37%	6%
2001	0.635	43%	11%	44%	3%
2002	0.570	54%	11%	34%	1%
2003	0.579	46%	10%	43%	1%
2004	0.636	50%	13%	34%	4%
2005	0.361	64%	14%	20%	2%
2006	0.491	62%	8%	28%	2%
2007	0.347	60%	14%	23%	3%
2008	0.408	67%	20%	20%	3%
2009	0.345	56%	19%	4%	21%
2010	0.326	78%	17%	2%	4%
2011	0.285	57%	9%	32%	2%
2012	0.454	33%	7%	57%	3%
2013	0.538	47%	17%	33%	3%
2014	0.370	57%	17%	22%	4%
2015	0.516	35%	11%	54%	1%
2016	0.273	48%	15%	33%	4%
2017	0.413	61%	16%	22%	1%
2018	0.375	39%	23%	33%	4%
2019	0.329	46%	18%	34%	2%
2020	0.104	55%	6%	8%	31%
2021	0.406	61%	15%	20%	4%
2022	0.432	36%	12%	50%	1%

Table 5. The proportion of VIF, LMG and VIM values of each variable in 2020.

	R ²	2 m Temperature	10 m Wind Speed	Sea Level Pressure	Sea Ice Motion
VIF		1.04	1.50	1.01	1.51
VIM	0.765	31%	19%	25%	25%
LMG	0.104	55%	6%	8%	31%

Combined with the actual case to analyze: In June 2020, the average temperature in the Arctic Siberian region was 10 °C higher than normal, sparking massive wildfires in the region [44]. The year 2020 is also the second year in a row that Siberia and the Arctic Circle have seen massive wildfires. According to the World Meteorological Organization, 2020 is one of the three hottest years on record (the other two are 2016 and 2019). High temperatures in Siberia have led to the melting of large areas of sea ice in the Arctic Circle, contributing to the extension of navigation in 2020.

Due to the occurrence of extreme high-temperature events in 2020, the importance of the LMG and VIM for the 2 m temperature accounted for 55% and 31%, respectively. Higher temperatures lead to the melting of sea ice, which in turn leads to increased sea ice movement, which is the second most important of the two models. It also shows that the LMG is better at identifying extreme events, and RF is better at identifying the combined effects of all variables on sea ice intensity.

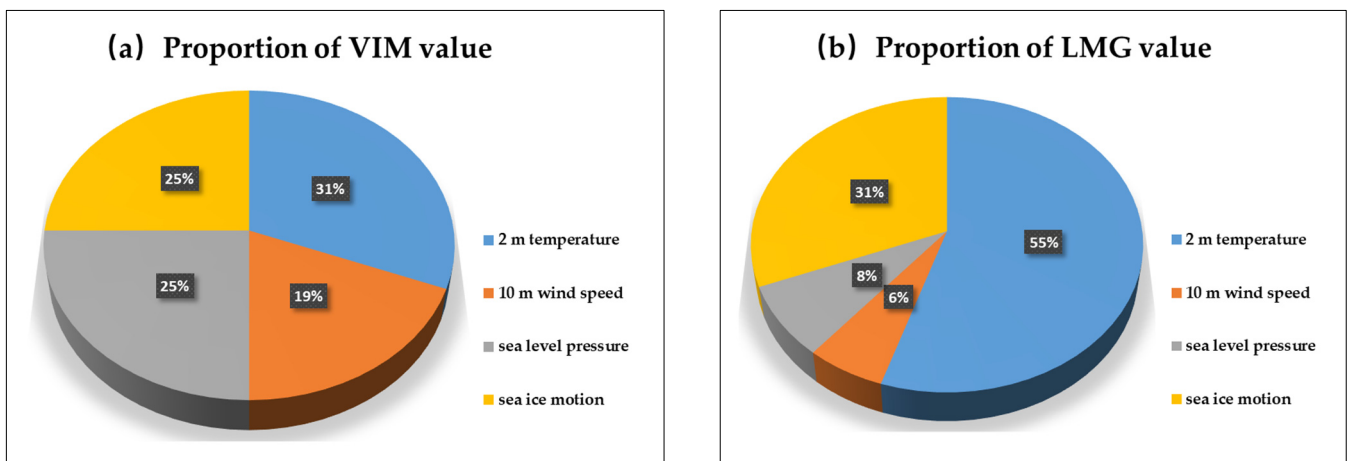


Figure 4. (a) Pie chart distribution of VIM value of each variable in 2020. (b) Pie chart distribution of LMG value of each variable in 2020.

3.4.1. Effect of 2 m Temperature on SIC

In Table 2, the proportion of VIM values for the 2 m air temperature varies from 39% to 51%, showing a relatively stable trend. In Table 3, the proportion of LMG values of the 2 m air temperature fluctuates between 33% and 78%, showing a large range in variation, indicating that its contribution to the overall model fluctuates greatly.

According to Figure 5, from July to September 2020, the areas with temperatures ranging from 0–10 °C gradually decreased. July and August experienced relatively higher temperatures, with some regions exceeding 10 °C, which accelerated the melting of sea ice and extended the navigable period. However, entering October, temperatures significantly dropped, with most areas outside the ends of the NEP falling below 0 °C. This cooling trend led to the reformation of sea ice, adversely affecting navigational conditions.

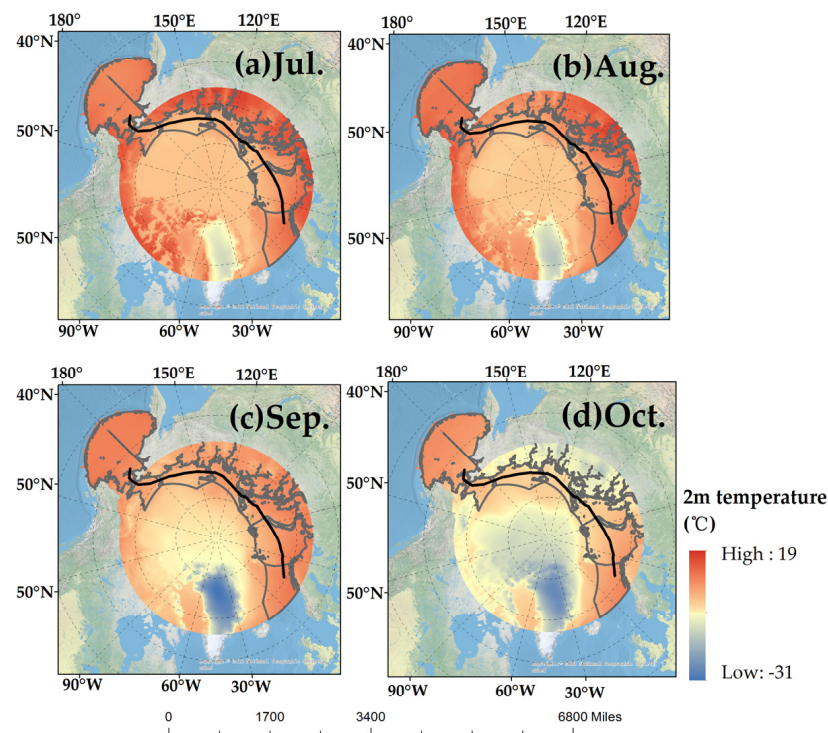


Figure 5. Monthly average distribution of 2 m temperature in 2020. The black line segment is the Northeast Passage, and the area composed of the gray line segment is the sea area involved in the Northeast Passage.

In the LMG, the relative importance coefficient of the 2 m temperature is highest, indicating its greatest contribution to the SIC. This suggests that the temperature is the primary factor driving changes in the SIC, particularly during summer and early autumn when higher temperatures accelerate sea ice melting. Similarly, in the RF importance evaluation, the 2 m temperature is also ranked as the most significant predictor variable. This further confirms the dominant role of temperature in influencing the SIC. RF captures the complex nonlinear relationship between temperature and the SIC, revealing that the temperature directly affects sea ice melting and formation and interacts with other climate factors to have a comprehensive impact on the SIC. For instance, the interaction between sea surface temperature and atmospheric temperature has a significant effect on the melting and freezing rate of sea ice, and its interaction with the 2 m temperature may lead to a dynamic change in sea ice cover [45]. The combined effect of ocean salinity and temperature will affect the structure and stability of sea ice. The rise in temperature may aggravate the melting of sea ice, thus affecting the density of sea ice [46].

3.4.2. Effect of 10 m Wind Speed on SIC

In Table 2, the proportion of VIM values of the 10 m wind speed varies from 14% to 29%, showing a certain fluctuation. In Table 3, the proportion of the LMG value of the 10 m wind speed fluctuates between 5% and 26%, which also shows a relatively large range of variation.

Based on Figure 6, in July, the 10 m wind speed primarily ranged between 4–6 m/s, indicating relatively low and stable wind speeds over a large area. In August, there was a slight increase in the 10 m wind speed, with the area of 4–6 m/s gradually decreasing and the 6–8 m/s range becoming more prominent, showing a gradual strengthening of winds. This trend continued into September, where the average monthly wind speed centered mainly around 6–8 m/s, with the 4–6 m/s range further diminishing, indicating a more pronounced trend of increasing wind speed. By October, the 10 m wind speed further intensified, reaching ranges of 8–11 m/s, indicating the highest wind speed of the year. From July to October 2020, there was a gradual increase in the 10 m wind speed, transitioning from the 4–6 m/s range to the 8–11 m/s range. The increase in wind speed not only significantly influences the SIC but also provides favorable conditions for navigational access along the passage.

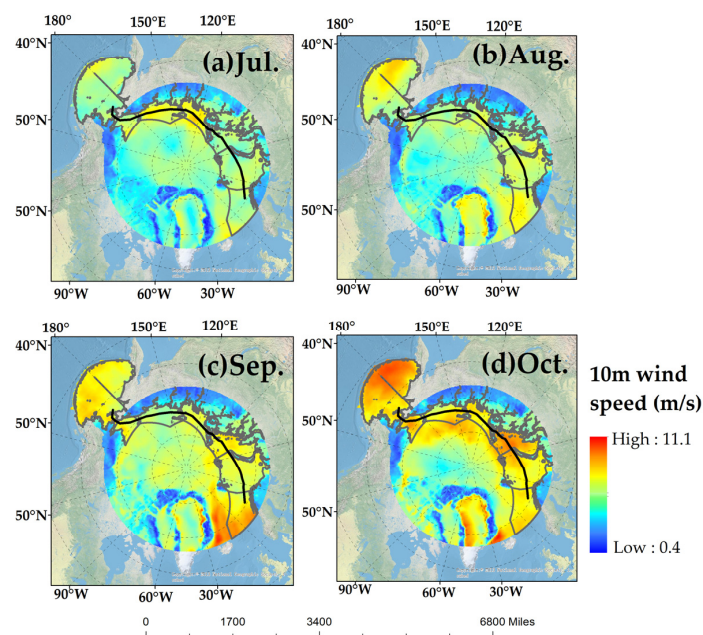


Figure 6. Monthly average distribution map of 10 m wind speed in 2020. The black line segment is the Northeast Passage, and the area composed of the gray line segment is the sea area involved in the Northeast Passage.

In 2020, recorded as the year with the longest navigable period in the NEP, this characteristic aligns with the trend of 10 m wind speed changes. Higher wind speeds facilitate the fracturing and drifting of sea ice, accelerating its melting rate and thus extending the navigable period. At the same time, higher wind speeds may also affect the distribution and density of sea ice, thereby influencing navigational conditions along the passage. Understanding these wind speed variations is crucial for comprehending the dynamics of sea ice and assessing the potential for future navigability of the NEP.

3.4.3. Effect of Sea Level Pressure on SIC

In Table 2, the proportion of the VIM values of sea level pressure fluctuates between 11% and 29%. In Table 3, the LMG ratio of sea level pressure varies from 1% to 31%.

Based on Figure 7, in July and August, a high-pressure system formed in the waters of the NEP, with pressures ranging from 1018–1023 hPa, and the extent of this high-pressure system decreased over time. Meanwhile, in other areas, the atmospheric pressure ranged between 1006–1018 hPa. The presence of a high-pressure system typically accompanies stable weather conditions and lower wind speeds, which helps reduce the motion and dispersion of sea ice, thereby affecting its distribution and density. Moving into September and October, a low-pressure center appeared in the area, with pressures ranging from 1000–1006 hPa, and like the high-pressure system, the extent of this low-pressure system decreased over time, gradually moving towards the central part of the NEP. Atmospheric pressures in other areas remained between 1006–1018 hPa. Low-pressure systems typically bring more storms and higher wind speeds, which can lead to ice fracturing and a faster drift.

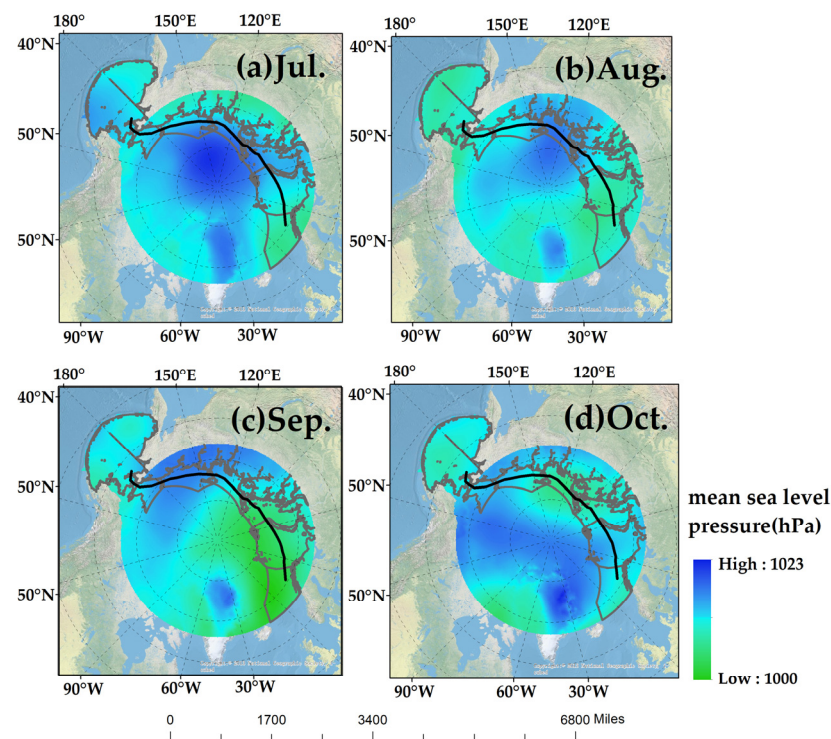


Figure 7. Monthly average distribution of sea level pressure in 2020. The black line segment is the Northeast Passage, and the area composed of the gray line segment is the sea area involved in the Northeast Passage.

In 2020, which was characterized by a longer navigable period, the spatiotemporal variations of high and low-pressure systems had significant impacts on the navigability of the Northeast Passage. The high-pressure centers in July and August provided more stable weather conditions, likely contributing to keeping the passage open and reducing the risks of ice restructuring and drift. Conversely, the low-pressure systems in September

and October, while potentially increasing the risks of storms and ice fracturing, may have accelerated the clearing of ice blocks, providing a brief window for navigation.

3.4.4. Effect of Sea Ice Motion on SIC

In Table 2, the proportion of the VIM values of sea ice drift fluctuates between 18% and 29%, showing a relatively large range of variation. In Table 3, the LMG ratio of sea ice drift varies from 20% to 57% and also shows a large fluctuation range.

Based on Figure 8, in July 2020, the sea ice motion range was the largest, with speeds ranging from 1–12 cm/s. During this period, the extensive motion facilitated rapid dispersion and reduced the local SIC, thereby positively impacting navigational conditions. As time progressed from August to October, the sea ice motion range gradually decreased each month, with speeds ranging from 1–8 cm/s. This reduction in motion range may have led to ice reaggregation, gradually increasing the SIC, but overall, it remained at relatively low levels, supporting the conclusion that 2020 experienced a longer navigable period.

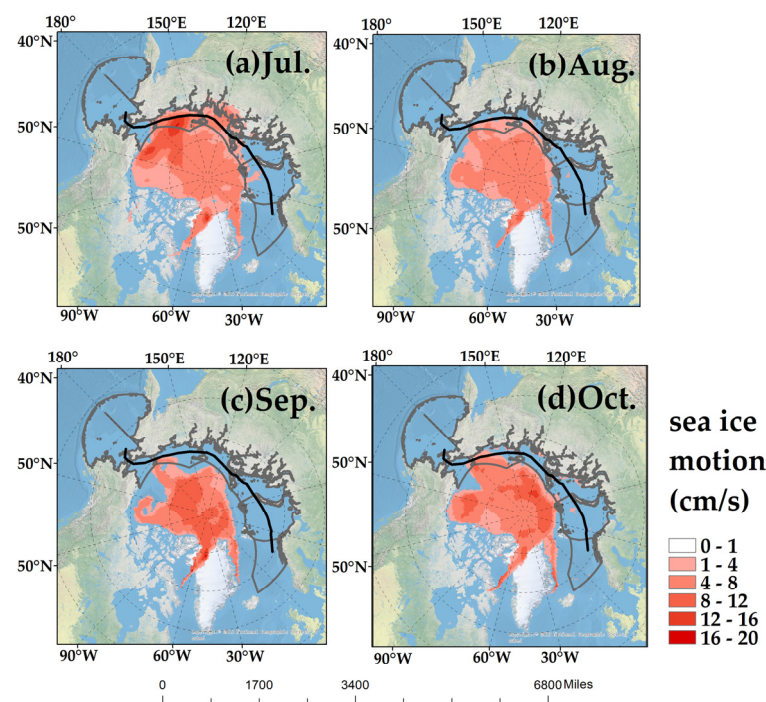


Figure 8. Monthly average distribution map of sea ice motion in 2020. The black line segment is the Northeast Passage, and the area composed of the gray line segment is the sea area involved in the Northeast Passage.

In the LMG, the sea ice motion's importance percentage is 8%, indicating that under the assumption of a linear relationship between independent and dependent variables, the impact of sea ice motion on the SIC is relatively small. However, in RF, the importance percentage of sea ice motion significantly increases to 25%. This suggests that when considering more complex nonlinear relationships and interactions, the influence of sea ice motion on the SIC becomes more significant. This difference likely reflects the complex nonlinear relationships between sea ice motion and other climatic factors.

4. Discussion

In the process of exploring the dynamic changes in the SIC in NEP, researchers have made continuous efforts to fill the gaps in existing research. In previous studies, the dynamic changes in the SIC and its key influencing factors during NEP navigation were not studied in combination with specific research methods. This study aims to fill this research gap and explore the characteristics and influencing factors of the SIC changes through MLR and RF.

The results of this study show that there is a 25-year navigable window from 1979 to 2022. The start date of these windows is roughly between late July and early August, and the end date is between early October and mid-October. The duration of these windows varies greatly. RF and MLR show that the 2 m temperature is the main variable affecting the SIC, and the sea ice movement, sea level pressure and 10 m wind speed also play a certain role in a specific period of time.

Ji et al. [47] used the Arctic SIC data and meteorological elements (temperature and wind speed) to analyze the factors affecting the navigation of the Arctic waterway, but this was only based on the monthly average data of meteorological elements. Gui et al. [15] studied how the change in sea ice motion affects the navigation capacity of the Arctic Passage, focusing on the role of dynamic factors such as sea ice drift and wind speed, but the conclusion obtained in this paper is that the 2 m temperature has the greatest impact on the SIC. Ruibo et al. [48] emphasized how meteorological anomalies, such as changes in wind speed and temperature, affect the SIC in the Arctic Ocean. It uses atmospheric reanalysis data to track sea ice movements and conditions, providing insights into how short-term atmospheric conditions affect sea ice intensity. This paper aims to study the changes in the SIC during the navigation period.

However, this paper also has some limitations. The limitation of the data's time range, the instability of early navigation time and the accuracy of the model may have a certain impact on the results. Future research can further expand the time and space range of data, improve the accuracy of the model and understand the dynamic changes in the SIC more comprehensively. More advanced statistical methods can be further combined to improve the accuracy and reliability of the model.

5. Conclusions

This paper is based on the SICs from 1979–2022, along with thermodynamic and dynamic-related factors. Using MLR and RF, it analyzes the navigational windows of the NEP and explores key factors influencing the dynamic changes in the SIC. Below are the research conclusions:

- (1) From 1979 to 1989, the NEP was consistently blocked by high-density sea ice, resulting in no navigable windows during this period. From 1990 to 2022, there were 25 years with navigable windows, characterized by significant variability in the navigational periods. The longest navigational period was 73 days in 1995, while the shortest was only 8 days in 1997. Between 2001 and 2011, with the exception of 2001, 2003 and 2004, the navigable windows generally lasted from 30 to 90 days. From 2012 to 2022, the variability in navigational periods decreased, with most years having navigable periods ranging from 50 to 110 days.
- (2) In order to compare the advantages of linear and nonlinear models in evaluating the importance of sailing cycles, MLR and RF were used in this study. The results showed that the explanatory power (R^2) of RF was significantly higher than that of MLR. In addition, combined with the actual navigation situation in 2020, it shows that the LMG is better at identifying extreme events, and RF is better at identifying the comprehensive impact of all variables on the sea ice concentration.
- (3) This research, combining traditional statistical methods and machine learning techniques, reveals the dynamic changes in the SIC during navigation in the Northeast Passage and its driving factors. The study identifies that the 2 m temperature is the primary variable influencing the SIC, while the sea ice movement, sea level pressure and 10 m wind speed also play a role during specific periods. By quantifying the relative importance of these factors, researchers can more accurately determine the factors that decisively impact navigability under specific conditions and time frames, thereby optimizing shipping decisions.

Subsequently, the 2 m air temperature can be used as a predictor to predict the sea ice concentration. Analyzing the dynamic changes in the sea ice concentration and its related factors in the future can help navigation planners understand the navigable window of the Northeast Passage more accurately so as to optimize route planning.

While the temperature has been identified as a primary influencing factor across different years, its specific impact may vary depending on the year and season. The impact of sea ice drift on the SIC during navigation is also worthy of attention. Although the influence of sea ice drift is relatively small in the linear model, its importance is improved in the nonlinear model, especially during navigation. This suggests that we need to pay more attention to the impact of sea ice drift when predicting navigation conditions and SIC changes, and combine other meteorological factors for comprehensive analysis.

Author Contributions: Conceptualization, Y.H.; Methodology, Y.L. (Yanhua Liu); Validation, D.F. and Y.L. (Yongheng Li); Formal analysis, F.J. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the National Natural Science Foundation of China (No. 4197060184).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data on Arctic sea ice concentrations were obtained from the National Snow and Ice Data Center (<https://nsidc.org/>, accessed on 15 February 2023).

Conflicts of Interest: Author Duxian Feng was employed by the company CPECC Xinjiang Petroleum Engineering Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wang, K.; Zhang, Y.; Chen, C.; Song, S.; Chen, Y. Impacts of Arctic Sea Fog on the Change of Route Planning and Navigational Efficiency in the Northeast Passage during the First Two Decades of the 21st Century. *J. Mar. Sci. Eng.* **2023**, *11*, 2149. [CrossRef]
2. Wu, A.; Che, T.; Xu, Q.; Chen, J.; Zhu, X. Assessing the economic viability of the Arctic Northeast Passage from 2021 to 2065. *Int. J. Digit. Earth* **2024**, *17*, 2323182. [CrossRef]
3. Geng, J.; Guan, L.; Wu, F.; Wang, H. Analysis on the Variation of Arctic Sea Ice Using Satellite Data. *J. Ocean. Technol.* **2014**, *33*, 8–13. (In Chinese)
4. Ma, L.; Li, Z.; Chen, G.; Li, Y. Research on the navigability of the arctic northeast route based on sea ice conditions during the passage of M/V YONG SHENG. *Chin. J. Polar Res.* **2018**, *30*, 173–185. (In Chinese)
5. Li, X.; Dong, C. Estimating Sea Ice Concentration from Microwave Radiometric Data for Arctic Summer Conditions Using Machine Learning. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4301018. [CrossRef]
6. Khon, V.C.; Mokhov, I.I.; Semenov, V.A. Transit navigation through Northern Sea Route from satellite data and CMIP5 simulations. *Environ. Res. Lett.* **2017**, *12*, 02400. [CrossRef]
7. Kibanova, O.V.; Eliseev, A.V.; Mokhov, I.I.; Khon, V.C. Bayes Estimates of Variations of the Duration of the Navigation Period Along the Northern Sea Route in the XXI Century from Simulations with Ensemble of Climate Models. In *Physical and Mathematical Modeling of Earth and Environment Processes (2018), Proceedings of the 4th International Scientific School for Young Scientists, Ishlinskii Institute for Problems in Mechanics of Russian Academy of Sciences, Moscow, Russia, 1–3 November 2018*; Karev, V., Klimov, D., Pokazeev, K., Eds.; Springer Proceedings in Earth and Environmental Sciences; Springer: Cham, Switzerland, 2019. [CrossRef]
8. Ivanov, V. Arctic Sea Ice Loss Enhances the Oceanic Contribution to Climate Change. *Atmosphere* **2023**, *14*, 409. [CrossRef]
9. Stroeve, J.; Holland, M.M.; Meier, W.; Scambos, T.; Serreze, M. Arctic sea ice decline: Faster than forecast. *Geophys. Res. Lett.* **2007**, *34*, L09501. [CrossRef]
10. Lee, H.J.; Kwon, M.O.; Yeh, S.W.; Kwon, Y.O.; Park, W.; Park, J.-H.; Kim, Y.H.; Alexander, M.A. Impact of poleward moisture transport from the North Pacific on the acceleration of sea ice loss in the Arctic since 2002. *J. Clim.* **2017**, *30*, 6757–6769. [CrossRef]
11. Li, Y.; Jiang, Z.; Yao, Y.; Ding, M.; Zhang, L. Two distinct declining trend of autumn Arctic sea ice concentration before and after 2002. *Environ. Res. Commun.* **2024**, *6*, 035018. [CrossRef]
12. Min, C.; Zhou, X.; Luo, H.; Yang, Y.; Wang, Y.; Zhang, J.; Yang, Q. Toward Quantifying the Increasing Accessibility of the Arctic Northeast Passage in the Past Four Decades. *Adv. Atmos. Sci.* **2023**, *40*, 2378–2390. [CrossRef]
13. Xiu, Y.; Wang, Q.; Li, Z. Estimating spatial distributions of design air temperatures for ships and offshore structures in the Arctic Ocean. *Polar Sci.* **2022**, *34*, 100875. [CrossRef]
14. Uhlíková, T.; Vihma, T.; Karpechko, A.Y.; Uotila, P. Effects of Arctic Sea-Ice Concentration on Turbulent Surface Fluxes in Four Atmospheric Reanalyses. *Cryosphere* **2024**, *18*, 957–976. [CrossRef]

15. Gui, D.; Pang, X.; Lei, R.; Zhao, X.; Wang, J. Changes in sea ice kinematics in the Arctic outflow region and their associations with Arctic Northeast Passage accessibility. *Acta Oceanol. Sin.* **2019**, *38*, 101–110. [\[CrossRef\]](#)
16. Chu, J.; Xiao, X. Benefits evaluation of the Northeast Passage based on grey relational degree of discrete Z-numbers. *Inf. Sci.* **2023**, *626*, 607–625. [\[CrossRef\]](#)
17. Liu, M.; Hu, S.; Ge, Y.; Heuvelink, G.B.; Ren, Z.; Huang, X. Using multiple linear regression and random forests to identify spatial poverty determinants in rural China. *Spat. Stat.* **2021**, *42*, 100461. [\[CrossRef\]](#)
18. IEEE. *Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, Harbin, China, 25–17 December 2020; IEEE: Piscataway, NJ, USA, 2020; p. 1. [\[CrossRef\]](#)
19. Grömping, U. Variable importance in regression models. *Wiley Interdiscip. Rev. Comput. Stat.* **2015**, *7*, 137–152. [\[CrossRef\]](#)
20. Wei, P.; Lu, Z.; Song, J. Variable importance analysis: A comprehensive review. *Reliab. Eng. Syst. Saf.* **2015**, *142*, 399–432. [\[CrossRef\]](#)
21. Grömping, U. Estimators of Relative Importance in Linear Regression Based on Variance Decomposition. *Am. Stat.* **2007**, *61*, 139–147. [\[CrossRef\]](#)
22. Lai, J.; Zou, Y.; Zhang, J.; Peres-Neto, P.R. Generalizing hierarchical and variation partitioning in multiple regression and canonical analyses using the rdacca. hp R package. *Methods Ecol. Evol.* **2022**, *13*, 782–788. [\[CrossRef\]](#)
23. Lindeman, R.H.; Merenda, P.F.; Gold, R.Z. *Introduction to Bivariate and Multivariate Analysis*; Scott Foresman: Glenview, IL, USA, 1980.
24. Budescu, D.V. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychol. Bull.* **1993**, *114*, 542–551. [\[CrossRef\]](#)
25. Feldman, B.E. Relative Importance and Value. *SSRN Electron. J.* **2005**. [\[CrossRef\]](#)
26. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
27. Pahlavan, H.A.; Zahraie, B.; Nasser, M. Improvement of multiple linear regression method for statistical downscaling of monthly precipitation. *Int. J. Environ. Sci. Technol.* **2018**, *15*, 1897–1912. [\[CrossRef\]](#)
28. Lin, H.Y.; Li, C.F.; Chen, T.Y.; Hsieh, C.F.; Wang, G.; Wang, T.; Hu, J.M. Climate-Based approach for modeling the distribution of montane forest vegetation in Taiwan. *Appl. Veg. Sci.* **2020**, *23*, 239–253. [\[CrossRef\]](#)
29. Oukawa, G.Y.; Krel, P.; Targino, A.C. Fine-Scale modeling of the urban heat island: A comparison of multiple linear regression and random forest approaches. *Sci. Total Environ.* **2022**, *815*, 152836. [\[CrossRef\]](#)
30. Guo, W.; Ni, X.; Mu, Y.; Liu, T.; Zhang, J. Detection and Attribution of Alpine Inland Lake Changes by Using Random Forest Algorithm. *Remote Sens.* **2023**, *15*, 1144. [\[CrossRef\]](#)
31. Liu, S.; Zou, B.; Shi, L.J.; Cui, Y. Polar SIC inversion method based on FY-3C microwave radiometer data. *Oceanogr. J.* **2020**, *42*, 113–122. (In Chinese)
32. Ke, C.Q.; Wang, M.M. Seasonal and interannual variation characteristics of Arctic sea ice thickness and volume from 2010 to 2017 based on CryoSat-2 data. *Chin. J. Oceanogr.* **2018**, *40*, e2021GL097448.
33. Chen, S.Y.; Cao, Y.F.; Hui, F.M.; Cheng, X. Spatio-temporal changes in the navigation capacity of the Arctic Northeast Passage in the autumn of 2010–2017 based on remote sensing observations. *Sci. Bull.* **2019**, *64*, 1515–1525. (In Chinese)
34. Shibata, H.; Izumiyama, K.; Tateyama, K.; Enomoto, H.; Takahashi, S. Sea-ice coverage variability on the Northern Sea Routes, 1980–2011. *Ann. Glaciol.* **2013**, *54*, 139–148. [\[CrossRef\]](#)
35. Fisher, R.A. 012: A Mathematical Examination of the Methods of Determining the Accuracy of an Observation by the Mean Error, and by the Mean Square Error. *Mon. Not. R. Astron. Soc.* **1920**, *80*, 758–770. [\[CrossRef\]](#)
36. Naimi, B.; Hamm, N.A.S.; Groen, T.A.; Skidmore, A.K.; Toxopeus, A.G. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **2014**, *37*, 191–203. [\[CrossRef\]](#)
37. Johnson, J.W.; Lebreton, J.M. History and Use of Relative Importance Indices in Organizational Research. *Organ. Res. Methods* **2016**, *7*, 238–257. [\[CrossRef\]](#)
38. Kruskal, W. Relative importance by averaging over orderings. *Am. Stat.* **1987**, *41*, 6–10. [\[CrossRef\]](#)
39. Grömping, U. Relative importance for linear regression in R: The package relaimpo. *J. Stat. Softw.* **2006**, *17*, 1–27. [\[CrossRef\]](#)
40. Liaw, A.; Wiener, M. Classification and regression by random Forest. *R. News* **2002**, *2*, 18–22.
41. Wright, M.N.; Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *arXiv* **2015**, arXiv:1508.04409. [\[CrossRef\]](#)
42. Strobl, C.; Boulesteix, A.L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinf.* **2007**, *8*, 25. [\[CrossRef\]](#)
43. Genuer, R.; Poggi, J.M.; Tuleau-Malot, C. Variable selection using Random Forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [\[CrossRef\]](#)
44. Shikhov, A.; Antokhina, O.; Gochakov, A.; Azhigov, I.; Blinov, D.; Gazimovet, T. Severe convective outbreak in Siberia in May 2020: Event analysis and high-resolution simulation. *Atmos. Res.* **2024**, *298*, 107139. [\[CrossRef\]](#)
45. Yuan, Y.; Li, H.; Sun, B.; He, S.; Li, F.; Li, H. Joint influence of the North Atlantic sea surface temperature and the Barents sea-ice concentration on the dipole pattern of Eurasian surface air temperature in March. *Atmos. Res.* **2024**, *305*, 107446. [\[CrossRef\]](#)
46. Yu, L.S.; He, H.; Leng, H.; Chen, X. Interannual Variation of Summer Sea Surface Salinity in the Dotson–Getz Trough, West Antarctica. *J. Mar. Sci. Eng.* **2024**, *12*, 1067. [\[CrossRef\]](#)

47. Ji, M.; Liu, G.; He, Y.; Li, Y.; Li, T. Analysis of Sea Ice Timing and Navigability along the Arctic Northeast Passage from 2000 to 2019. *J. Mar. Sci. Eng.* **2021**, *9*, 728. [[CrossRef](#)]
48. Ruibo, L.; Fanyi, Z.; Mengxi, Z. Meteorological and sea ice anomalies in the western Arctic Ocean during the 2018–2019 ice season: A Lagrangian study. *Adv. Polar Sci.* **2022**, *33*, 204–219.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.