*Article*

# Statistical Modeling of Arctic Sea Ice Concentrations for Northern Sea Route Shipping

Da Wu [1,2,3] , Wuliu Tian [1,2,3,4,5,*] , Xiao Lang [6,*] , Wengang Mao [6] and Jinfen Zhang [1,2,3]

1    State Key Laboratory of Maritime Technology and Safety, Wuhan University of Technology, Wuhan 430070, China
2    National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan 430070, China
3    Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430070, China
4    Guangxi Key Laboratory of Ocean Engineering Equipment and Technology, Beibu Gulf University, Qinzhou 535011, China
5    Key Laboratory of Beibu Gulf Offshore Engineering Equipment and Technology, Beibu Gulf University, Qinzhou 535011, China
6    Department of Mechanics and Maritime Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
*    Correspondence: tianwuliu@foxmail.com (W.T.); xiao.lang@chalmers.se (X.L.)

**Abstract:** The safe and efficient navigation of ships traversing the Northern Sea Route demands accurate information regarding sea ice concentration. However, the sea ice concentration forecasts employed to support such navigation are often flawed. To address this challenge, this study advances a statistical interpolation method aimed at reducing errors arising from traditional interpolation approaches. Additionally, this study introduces an autoregressive integrated moving average model, derived from ERA5 reanalysis data, for short-term sea ice concentration forecasts along the Northern Sea Route. The validity of the model has been confirmed through comparison with ensemble experiments from the Coupling Model Intercomparison Project Phase 5, yielding reliable outcomes. The route availability is assessed on the basis of the sea ice concentration forecasts, indicating that the route will be available in the upcoming years. The proposed statistical models are also shown the capacity to facilitate effective management of Arctic shipping along the Northern Sea Route.

**Keywords:** sea ice concentration; Arctic shipping; transit navigation windows; ARIMA model; time series analysis

## 1. Introduction

Sea Ice Concentration (SIC) is a paramount environmental indicator of climatic change in the Arctic region. In light of the fluctuations in SIC, the Northern Sea Route (NSR) has become the focus of growing interest in recent years, as detailed in [1]. Consequently, accurate forecasts of SIC are crucial for the safe and efficient navigation of ships along the NSR, as demonstrated in studies by [2–4]. Chen et al. [5] evaluated the decadal changes in sea ice parameters and Arctic navigability, which suggested open water ships will be able to cross the Northern Sea Route and Northwest Passage between August and October during the period from 2045 to 2055, with a maximum navigable percentage in September. Liu et al. [6] calculated the navigational window of some Arctic routes based on sea ice density data from 2006 to 2015 and gained a better understanding of the sea ice conditions in the Northwest Passage, and the findings showed that the shortest navigational window for each route is 69 days. Zhou et al. [7] integrated satellite sea ice thickness data into the feasibility assessment model of the Arctic navigation, analyzing the navigability of the Arctic routes. The study showed that the overall navigational conditions of the Northwest Passage are poor, and the navigational window is relatively short. To reduce uncertainty

in the prediction of the window period, Jun [8] researched the impact of radio detection observation on reducing uncertainty in Arctic sea ice prediction and identified issues regarding how to make forecast information convenient for crew and shipping companies to use. Li et al. [9] developed and verified two ship performance models which have been implemented in a voyage planning tool designed for summer Arctic operations of a cargo ship on the Northern Sea Route.

The SIC is typically recorded by satellite observations or by the observations of on-board master mariners. Researchers have utilized recorded data to formulate various models that attempt to discern the relationship between atmospheric anomalies and the sea ice regime. Stark et al. [10] incorporated the rheology mechanics of sea ice and its thickness distribution, ultimately positing that SIC is primarily governed by thermodynamics. Ye et al. [11] devised a scheme for reconstructing SIC using ice drift records, while Yu et al. [12] demonstrated that changes in SIC can be attributed to anomalous surface air temperatures and low-level wind fields. In a subsequent study [13], Ye et al. compared and appraised different methods for measuring sea ice type concentration. Mu et al. [14] proposed an ensemble-based Arctic Ice Ocean Prediction System, designed to facilitate forecasts of Arctic sea ice. However, physical models tend to be computationally intensive and require substantial amounts of data, and the resolution of available data may sometimes be insufficient to accurately model the geography along the NSR.

Furthermore, the statistical properties of sea ice concentration data are often neglected and under-investigated. As a statistical methodology, the autoregressive integrated moving average (ARIMA) model benefits from low computational requirements and high ease of use for time series analysis (TSA). Such techniques have been utilized in the context of oceanic environmental analysis, as demonstrated by Spanos [15], who compared an autoregressive (AR) model, a moving average (MA) model, and an autoregressive moving average (ARMA) model for simulating oceanic waves. Soares et al. [16] modeled significant wave heights through the use of an AR process that accounted for seasonality. Additionally, Agrawal and Deo [17] conducted a comparison between the ARMA model and an artificial neural network. These techniques have also been extended to a wide range of applications beyond oceanic waves, as seen in the works of Jiang et al. [18], who utilized the AR model for real-time prediction of ship motion, and Mao et al. [19], who applied the AR model for forecasting the speeds of container ships. The ARIMA model has been utilized in prior studies to examine the SIC. Piwowar and LeDrew [20] elucidated the autoregressive character of Arctic sea ice concentration. Ahn et al. [21] conducted a reanalysis of Arctic sea ice data through the ARIMA model, taking into account the changing temporal relationships between sea ice concentration and other climatic factors. Despite these studies, there remain limited instances where statistical models have been employed for the forecasting of Arctic sea ice concentration with a view toward the management of shipping along the Northern Sea Route.

This study proposes an ARIMA model to analyze and forecast the SIC along the NSR. A statistical interpolation approach is introduced to contract the SIC data obtained from the ERA5 [22] gridded reanalysis values into values along the NSR. The ARIMA model is trained and tested using reanalysis data from January 1980 to September 2022. Forecasts are generated for the period from 2022 to 2025 and validated through a comparison with ensemble experiments from the CMIP5 [23]. Based on this validation, an estimate for the 80% ice-free transit navigation window has been determined to serve as a guide for future NSR shipping. The results suggest that the proposed statistical methods are suitable and advantageous for the purpose of managing and planning NSR shipping.

This paper is organized as follows: Section 2 presents the statistical interpolation, and the ARIMA model is described in Section 3. In Section 4, the method is implemented as an example, and the results are discussed, while the conclusions are given in Section 5.

## 2. Interpolation of Sea Ice Concentrations

A typical NSR is depicted by the red trajectory in Figure 1. In accordance with the navigation plan of the ship that traversed this route, the NSR has been divided into eight segments, and consequently, eight sub-regions (depicted by green polygonal regions) have been established, each of which encompasses a single segment. In each sub-region, two longitudinal boundaries have been demarcated that define the starting and ending longitudes that the vessel traversed in accordance with the navigation plan; the northern boundary is situated 4 degrees north of the highest latitude reached on the route segment split in the section, while the southern boundary is located 4 degrees south of the lowest latitude reached on the route segment.
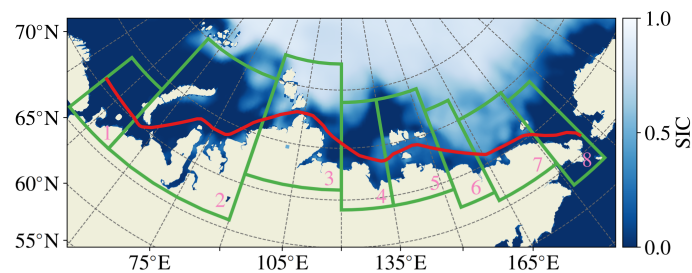


**Figure 1.** The NSR (red line) and representative sub-regions, as defined in the text, with sea ice conditions chosen from the ERA5 single-level monthly mean values for the date 1 July 2019.

The transit navigation window, a critical determinant of a shipping company's plans, is defined by the availability of the route. The acquisition of SIC data along the NSR from the ERA5 grid point values is indispensable in determining route availability. In the event that the grid resolution is adequate, interpolation from the nearest grid point values can yield acceptable SIC data along the route. However, such high resolution may not always be attainable, especially in areas of narrow straits. In such cases, interpolation necessitates a more comprehensive understanding of SIC in the entire sub-region, rather than being confined to the nearest grid point values. Additionally, when forecasting SICs for future shipping plans, the predicted high-resolution SICs often introduce significant uncertainties. In other words, it is easier to ascertain the presence of ice coverage along an extended route than to determine the SIC values at specific geographic locations. Consequently, the mean SIC of each route segment provides a more credible representation.

In Figure 2, the traditional method of interpolation via the nearest grid points is depicted on the left, while the proposed method, which leverages all grid values within the sub-region to enhance the interpolation, is presented on the right. A vast region provides significantly more information than a limited number of nearest grid points. For the purpose of statistical interpolation, the following features were selected from the gridded SIC values within a sub-region: total sea ice area, mean, variance, standard deviation, skewness, and kurtosis. A more extensive discourse on the statistical interpolation procedure is presented in the following section.
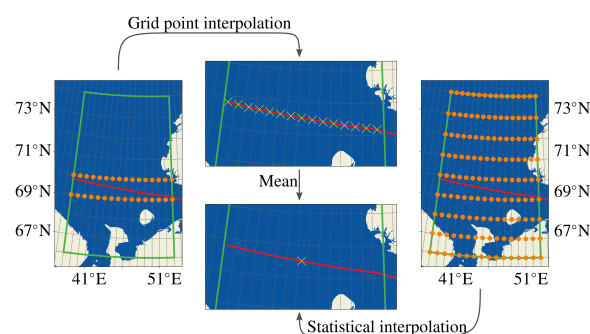


**Figure 2.** The grid point and statistical interpolations in sub-region No. 1.

*Statistical Interpolation of SIC*

For each sub-region, let $Y_t$, $X_{t,sia}$, $X_{t,mean}$, $X_{t,var}$, $X_{t,std}$, $X_{t,skew}$, and $X_{t,kurtosis}$ denote the mean SIC of the route segment, area covered by sea ice, mean, variance, standard deviation, skewness, and kurtosis of SIC in the sub-region at time $t$, respectively. Thus, the mean SIC of the route segment is assumed to be described by the linear model as follows:

$$
\begin{cases}
Y_1 &= \beta_0 + \beta_1 X_{1,sia} + \beta_2 X_{1,mean} + \beta_3 X_{1,var} + \beta_4 X_{1,std} + \beta_5 X_{1,skew} + \beta_6 X_{1,kurtosis} + \epsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{2,sia} + \beta_2 X_{2,mean} + \beta_3 X_{2,var} + \beta_4 X_{2,std} + \beta_5 X_{2,skew} + \beta_6 X_{2,kurtosis} + \epsilon_2 \\
\vdots \\
Y_N &= \beta_0 + \beta_1 X_{N,sia} + \beta_2 X_{N,mean} + \beta_3 X_{N,var} + \beta_4 X_{N,std} + \beta_5 X_{N,skew} + \beta_6 X_{N,kurtosis} + \epsilon_N
\end{cases}
\tag{1}
$$

and in matrix notation:

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} =
\begin{bmatrix}
1 & X_{1,sia} & X_{1,mean} & X_{1,var} & X_{1,std} & X_{1,skew} & X_{1,kurtosis} \\
1 & X_{2,sia} & X_{2,mean} & X_{2,var} & X_{2,std} & X_{2,skew} & X_{2,kurtosis} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & X_{N,sia} & X_{N,mean} & X_{N,var} & X_{N,std} & X_{N,skew} & X_{N,kurtosis}
\end{bmatrix}
\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{bmatrix} +
\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}.
\tag{2}
$$

Finally, the linear model can be expressed in a compact form as

$$
Y = X\beta + \varepsilon,
\tag{3}
$$

where $\beta$ represents coefficients to be estimated, $X$ is the design matrix, and $\varepsilon$ denotes errors. $\beta$ can be estimated by minimizing the sum of squared residuals (RSS),

$$
\begin{aligned}
\arg \min_{\beta_0, \cdots, \beta_6} \text{RSS}(\beta) &= \arg \min_{\beta_0, \cdots, \beta_6} \left( \varepsilon^\top \varepsilon \right) \\
&= \arg \min_{\beta_0, \cdots, \beta_6} \left( (Y - X\beta)^\top (Y - X\beta) \right) \\
&= \arg \min_{\beta_0, \cdots, \beta_6} \left( Y^\top Y - \beta^\top X^\top Y - Y^\top X\beta + \beta^\top X^\top X\beta \right),
\end{aligned}
\tag{4}
$$

which leads to

$$
\frac{\partial RSS(\widehat{\beta})}{\partial \widehat{\beta}} = -2X^\top Y + 2X^\top X\widehat{\beta} = 0.
\tag{5}
$$

Finally, the estimator $\widehat{\beta}$ can be expressed as

$$
\widehat{\beta} = \left( X^\top X \right)^{-1} X^\top Y,
\tag{6}
$$

under the assumption that $\text{rank}(X) = k + 1$ or

$$
\det\left( X^\top X \right) \neq 0.
\tag{7}
$$

Additionally, strict exogeneity tells that the conditional expectation of $\varepsilon$, given the design matrix $X$, is equal to zero, which implies that

$$
\widehat{Y} = \mathbb{E}(Y \mid X) = X\widehat{\beta}.
\tag{8}
$$

With Equation (8), a linear model has been developed to compute mean SICs of the NSR segments based on both the reanalysis SICs and their statistics.

## 3. Modeling of SIC and Time Series Analysis

The historical mean SICs of each NSR segment have been determined by the methods given in Section 2, forming a series of data points indexed in time order. This type of data can be described as a time series. In this section, we use the ARIMA model to analyze the SIC series.

### 3.1. Formulation of the ARIMA Model

The pre-requisite for a time series to be modeled by an ARIMA model is the demonstration of both stationarity and non-randomness. Stationarity greatly simplifies the complexity of the analysis and enhances the accuracy of estimation. Moreover, it helps to mitigate the risk of spurious regression [24]. However, a completely random series, such as white noise, is devoid of any correlations and therefore has no memory or influence on future values. That means that past information has no bearing on future information, and as such, it has no value for further analysis. Hence, before applying the ARIMA model, it is crucial to perform a stationarity test using the Augmented Dickey-Fuller (ADF) hypothesis test [25].

To represent the ARIMA model, we introduce the lag operator $B$ first. When a SIC series $y$ at time $t$ is multiplied by $B$ once, it means that time $t$ has lagged one step. For instance,

$$
\begin{aligned}
y_{t-1} &= B\ y_t, \\
y_{t-2} &= B^2 y_t, \\
&\ \ \vdots \\
y_{t-p} &= B^p y_t.
\end{aligned}
\tag{9}
$$

A stationary and non-random time series contains the relevant information extracted by an ARMA model. The information at the current time $t$ depends on the information in the past, which can be expressed mathematically as

$$
\begin{cases}
y_t = \underbrace{\varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p}}_{\text{auto-regressive (AR)}} + \underbrace{\epsilon_t - \vartheta_1 \epsilon_{t-1} - \cdots - \vartheta_q \epsilon_{t-q}}_{\text{moving average (MA)}}, \\
\quad = \varphi_1 B y_t + \cdots + \varphi_p B^p y_t + \epsilon_t - \vartheta_1 B \epsilon_t - \cdots - \vartheta_q B^q \epsilon_t, \\
\varphi_p \neq 0, \vartheta_q \neq 0, \\
\mathbb{E}(\epsilon_t) = 0, \mathbb{V}ar(\epsilon_t) = \sigma_\epsilon^2, \mathbb{E}(\epsilon_t \epsilon_s) = 0, s \neq t, \\
\mathbb{E}(y_s \epsilon_t) = 0, \forall s < t.
\end{cases}
\tag{10}
$$

Equation (10) implies that

1. the ARMA model can be written as ARMA $(p, q)$, which indicates that the highest order of the auto-regressive (AR) model is $p$, and the highest order of the moving average (MA) model is $q$;
2. the error $\epsilon_t$ is white noise; and
3. the error at the current time is independent of past $y$.

If the series is differenced to obtain the stationarity, and the series contains seasonal components, the ARMA model can be extended to an auto-regressive integrated moving average model, which is expressed in a compact form as

$$
\nabla^d \nabla_S^D y_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)} \varepsilon_t,
\tag{11}
$$

where

$$
\begin{aligned}
\nabla^d &= (1-B)^d, & \nabla_S^D &= (1-B^S)^D, \\
\Theta(B) &= 1 - \vartheta_1 B - \cdots - \vartheta_q B^q, & \Phi(B) &= 1 - \varphi_1 B - \cdots - \varphi_p B^p, \\
\Theta_S(B) &= 1 - \theta_1 B - \cdots - \theta_Q B^{QS}, & \Phi_S(B) &= 1 - \phi_1 B - \cdots - \phi_P B^{PS}.
\end{aligned}
\tag{12}
$$

The operator $\nabla^d$ is the ordinary difference component that represents the $d$-th differenced time series. The seasonal difference of order $D$ is defined as $\nabla_S^D$, where $S$ denotes the seasonal period, e.g., $S = 12$ for a yearly seasonal period. $\Theta(B)$ and $\Phi(B)$ are the nonseasonal autoregressive and moving average lag polynomials, respectively. The remaining seasonal information is then extracted by the seasonal autoregressive and moving average lag polynomials, $\Theta_S(B)$ and $\Phi_S(B)$, respectively. Equation (11) is denoted as ARIMA $(p, d, q) \times (P, D, Q, S)$ in short.

### 3.2. Model Identification and Forecast

For an ARIMA $(p, d, q) \times (P, D, Q, S)$ model, the parameter $d$ can be chosen by testing the stationarity of the $d$-th differenced series; $S$ can be obtained from the seasonal decomposition, e.g., the Census X-11 method [26]. The remaining orders $p$, $P$, $q$, and $Q$ can be determined from the properties of the AR($p$) and MA($q$) models, respectively.

Given an arbitrary AR($p$) model, from its partial autocorrelation function (PACF), we say that the AR($p$) model is "cutting off" at lag $p$. Consequently, $p$ and $P$ can be determined by this property from the PACF of series. Regarding an arbitrary MA($q$) model, its autocorrelation function (ACF) has the property of "cutting off" at $q$ lags, which can help to find the order of $q$ and $Q$. Once the orders have been determined, the parameters $\vartheta_i$, $\theta_i$, $\varphi_i$, and $\phi_i$ in Equation (11) are then estimated by using the maximum likelihood estimation (MLE) [27]. Finally, the minimum mean square error (MMSE) forecast [28] can be performed to produce a prediction.

## 4. Model Implementation and Validation

In this section, the statistical interpolation (described in Section 2) and the ARIMA model (described in Section 3) are applied and validated on data from ERA5 single-level monthly mean values for identifying the route availability. The workflow is illustrated in Figure 3. Sub-region No. 2 is chosen as an example and its context is described in detail.
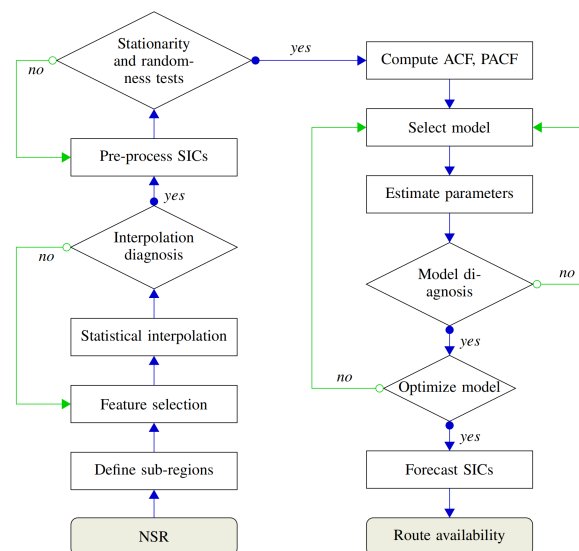


**Figure 3.** The workflow for implementing the statistical interpolation and ARIMA model for evaluating the route availability.

### 4.1. Statistical Interpolation in Sub-Regions

The results of the feature importance analysis suggest that the SIC metrics of area, mean, variance and standard deviation are of significant relevance to the statistical interpolation, as depicted in Figure 4. Conversely, skewness and kurtosis do not seem to have any significant impact. Additionally, a close examination of Pearson's coefficients, as illustrated in Figure 5, highlights strong linear correlations between the SIC's area and mean, as well as between its standard deviation and variance. In this regard, the sea ice area, $X_{t,sia}$,

and standard deviation, $X_{t,std}$, are selected out of the six statistics as features upon which to perform the interpolation. The results of the statistical regression are summarized in Table 1, and the SIC value can be expressed as

$$\widehat{Y}_t = 2.6803 X_{t,sia} - 0.2203 X_{t,std}. \tag{13}$$

In Table 1, the *t*-test suggests that the null hypothesis of $\boldsymbol{\beta = 0}$ can be rejected since *p*-values of $\beta_1$ and $\beta_2$ are equal to zero. Additionally, the $R^2$ value, which measures the proportion of variation in the dependent variables, is equal to 0.937 and also verifies that the statistical regression fits the reanalysis data well.

**Table 1.** Statistical interpolation results.

|  | Coeff. | Std. Err. | *t*-Statistic | *p*-Value |
|---|---|---|---|---|
| $\hat{\beta}_1$ | 2.6803 | 0.056 | 47.970 | 0.000 |
| $\hat{\beta}_2$ | −0.2203 | 0.025 | −8.899 | 0.000 |
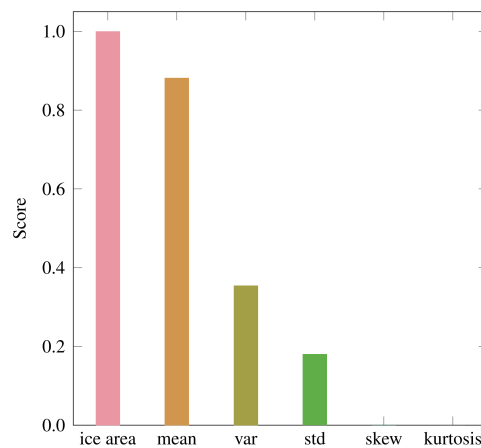


**Figure 4.** Feature ranking for statistical interpolation in sub-region No. 2.



**Figure 5.** The Pearson's coefficients of features in sub-region No. 2.

Following the implementation of statistical interpolation across all sub-regions, the correlation between the interpolated and reanalysis SIC values is depicted in Figure 6. This illustrates that most of the points are situated in close proximity to the diagonal line, suggesting that the linear model effectively captures the SIC along the NSR. A diagnostic examination of the residual is depicted in Figure 7, which demonstrates, through the Q-Q plot and histogram, that the residual of the interpolation conforms to a normal distribution as intended. The scatter plots of "residual vs. fitted values" and "residual vs. *X*" evince

randomness, thus affirming that the residual of the statistical interpolation constitutes white noise.
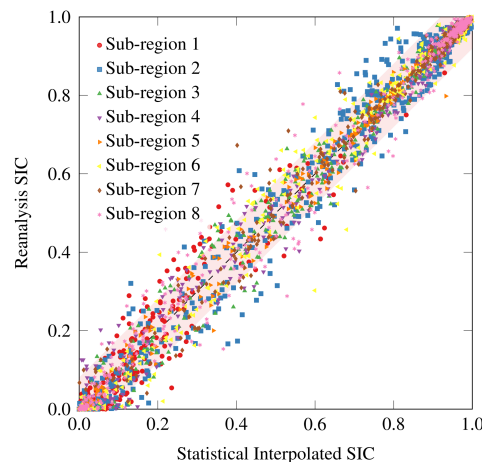


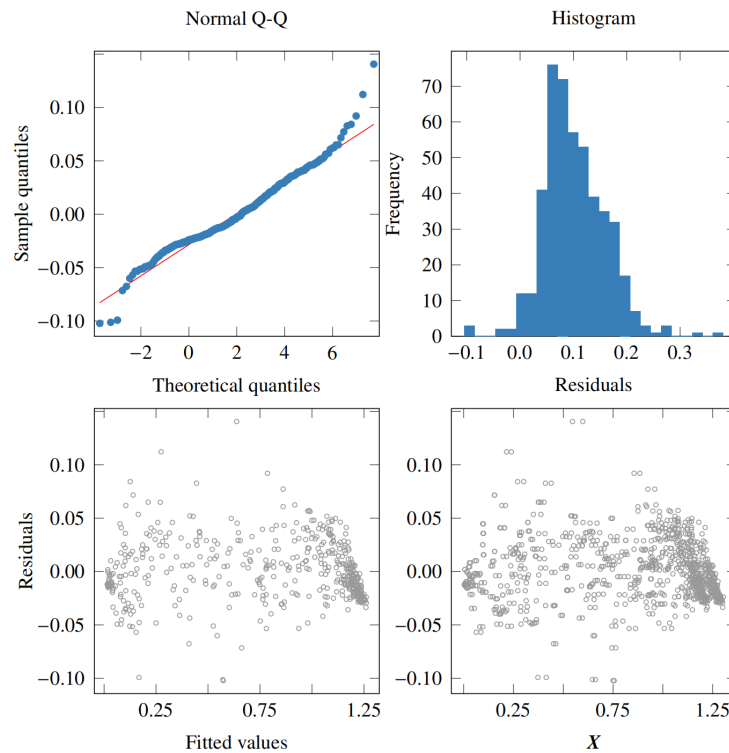**Figure 6.** The correlation between interpolated and reanalysis SICs along the NSR.



**Figure 7.** The residual diagnosis of the statistical interpolation in sub-region No. 2.

As previously noted in Section 2, the precision of the nearest grid point interpolation is strongly contingent upon the resolution of the grid. To evaluate the sensitivity of the statistical interpolation to grid size, the nearest grid point interpolation and statistical interpolation were performed with grid sizes ranging from 0.5° by 0.5° to 3.5° by 3.5°. The nearest grid point interpolation with a resolution of 0.25° by 0.25° served as the reference. As depicted in Figure 8, it is evident that as the grid size increases, the nearest grid point

interpolation progressively generates larger root-mean-squared errors (RMSEs, as indicated by Equation (14)).

$$e_i = \sqrt{\frac{\sum_{t=1}^{N}\left(\widehat{Y}_{t,m,i} - \widehat{Y}_{t,0.25,1}\right)^2}{N}},$$

$$\text{with } m \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5\},$$

$$i = 1, 2,$$

(14)

where $\widehat{Y}_t$ is the interpolated SIC defined in Equation (8). The nearest grid point interpolation and statistical interpolation are denoted as $i = 1$ and 2, respectively. The total number of timestamps is represented by $N$, and $m$ represents the grid sizes. In fact, when the grid resolution changes from 0.5° to 3.5°, the RMSEs increase by a factor of ten. However, the statistical interpolation exhibits a comparatively consistent RMSE trend, fluctuating around 0.05. While, at smaller grid sizes, the statistical interpolation yields higher errors than the nearest grid point interpolation. As the grid size exceeds 2 degrees, both interpolation methods produce comparable errors. When the grid size exceeds 3 degrees, the statistical interpolation proves to be significantly more accurate, yielding much lower errors in comparison to the nearest grid point interpolation. The insensitivity of the statistical interpolation to grid resolution, as demonstrated in Figure 8, expands its practicality and versatility, as it can be effectively employed with both high-resolution and low-resolution grids.
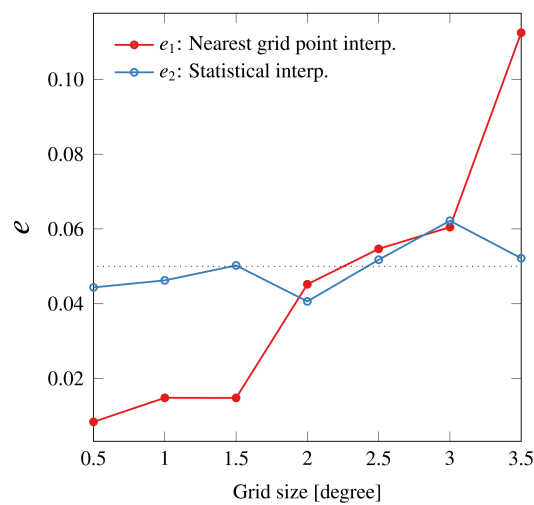


**Figure 8.** The errors of different grid sizes and methods.

In this study, we undertake a comparison of the proposed statistical interpolation method with that of kriging interpolation with various variogram models. In order to establish a baseline for comparison, we obtain the average SIC values within each region through the use of the nearest grid point interpolation method, utilizing the highest resolution reanalysis of SIC data. Subsequently, the average SIC values are determined through statistical interpolation and kriging interpolation, and the root means square errors are calculated and compared against the reference values. The results (in Table 2) reveal that the average SICs obtained through the statistical interpolation method exhibit lower RMSEs in all regions as compared to the kriging interpolation.

**Table 2.** The errors of statistical interpolation and kriging with different variogram models.

| Region | Linear | Power | Gaussian | Spherical | Exponential | Hole-Effect | Statistical |
|--------|--------|-------|----------|-----------|-------------|-------------|-------------|
| 1 | 0.139 | 0.110 | 0.101 | 0.117 | 0.093 | 0.076 | 0.043 |
| 2 | 0.178 | 0.143 | 0.182 | 0.199 | 0.199 | 0.157 | 0.060 |
| 3 | 0.071 | 0.076 | 0.100 | 0.101 | 0.081 | 0.102 | 0.036 |
| 4 | 0.221 | 0.202 | 0.206 | 0.196 | 0.201 | 0.134 | 0.035 |
| 5 | 0.157 | 0.154 | 0.159 | 0.162 | 0.154 | 0.171 | 0.022 |
| 6 | 0.178 | 0.181 | 0.189 | 0.159 | 0.151 | 0.135 | 0.040 |
| 7 | 0.124 | 0.118 | 0.199 | 0.152 | 0.130 | 0.143 | 0.029 |
| 8 | 0.307 | 0.222 | 0.262 | 0.276 | 0.239 | 0.135 | 0.048 |

Figure 9 depicts the October to December quarter mean reanalysis SIC (for reference), statistical interpolation, and kriging interpolation as an example. It is evident that the kriging interpolation produces the most pronounced deviation around the year 2010. Conversely, both the statistical and kriging interpolations produce results that are largely comparable, particularly after the year 2017. Hence, the proposed statistical interpolation can serve as an adequate alternative to the kriging interpolation. Furthermore, the statistical interpolation avoids the need for variogram model selection, yielding it more feasible to implement than the kriging interpolation.
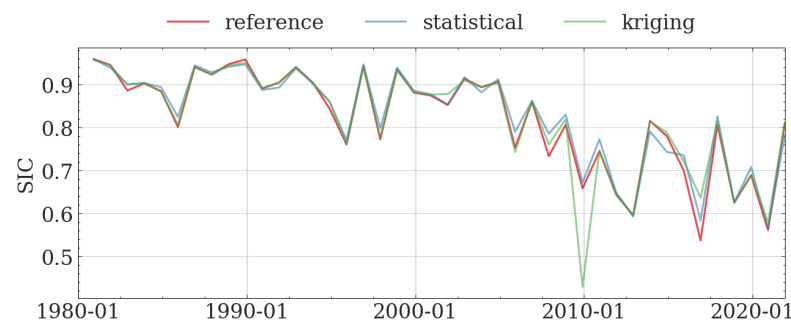


**Figure 9.** The comparison between kriging and statistical interpolation on SIC in October, November, and December.

### 4.2. Arima Modeling of SIC

The SICs along the given NSR obtained in Section 4.1 are plotted in Figure 10 as a time series. The time series contains a decreasing trend and a seasonal period of 12 months. To stabilize variances further, we preprocess SICs in two steps: (1) adding 1 to the original SICs and (2) performing a logarithmic transformation, i.e., $Y'_t = \text{Log}\left(\widehat{Y}_t + 1\right)$. The $Y'_t$ is then used to establish the ARIMA model. Therefore, the original SICs can be obtained from the back-transformation, $\widehat{Y}_t = e^{Y'_t} - 1$.

Following the flowchart (Figure 3), the tests of randomness and stationarity are required next. From Table 3, the Ljung-Box test suggests that the preprocessed SIC series in sub-region No. 2 contains relationships at lags of 1, 6, and 12 months. In other words, the series is non-random.

Next, the Augmented Dickey-Fuller Test (ADF) tests the null hypothesis that a series is not stationary. Before carrying out the test, a seasonal difference, $\nabla^1_{12} Y'_t$ (Equation (11)) is taken for the series. The results in Table 4 show that the p-value is smaller than 0.05, indicating that the seasonally differenced series are stationary. The statistic value is smaller than the critical values of 5% and 10%, and very close to 1%, which also supports the stationarity of the series.
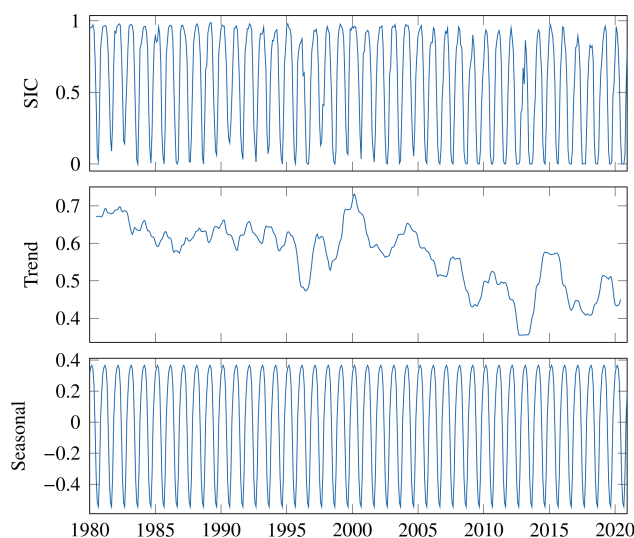
**Figure 10.** The SICs in sub-region No. 2 and the decomposed trend and seasonal components.

**Table 3.** Ljung-Box white noise test of SICs in sub-region No. 2.

| Lags | Ljung-Box Statistic | *p*-Value |
|---|---|---|
| 1 | 286.935 | 0.0 |
| 6 | 807.854 | 0.0 |
| 12 | 1746.735 | 0.0 |

**Table 4.** Augmented Dickey-Fuller test of seasonally differenced SICs in sub-region No. 2.

| Item | Value |
|---|---|
| *p*-value | 0.019 |
| test statistic | −3.208 |
| critical value (1%) | −3.446 |
| critical value (5%) | −2.868 |
| critical value (10%) | −2.570 |

In the ARIMA $(p, d, q) \times (P, D, Q, S)$ model, orders $d = 0$, $D = 1$, and $S = 12$ have been confirmed. The orders $p$, $q$, $P$, and $Q$ can be determined from the autocorrelation function (ACF) and the partial autocorrelation function (PACF) in Figure 11, the blue shaded regions represent the range of zero plus/minus two standard errors. When values fall within the shaded region, the ACF and PACF are considered to be zero. The lags of the red dots provide information to determine the orders of the ARIMA model. Within the first seasonal period, i.e., lag $h \in [0, 12]$, the ACF cuts off at lags 3, 4, or 5, whereas the PACF cuts off at lag 2. In this regard, $p$ is equal to 2, and $q$ can be either 3, 4, or 5. Furthermore, the ACF has a peak at lag 12 and decays very rapidly at other multiples of the seasonal period, e.g., $h = 12k$, for $k > 1$. Therefore, $Q$ is equal to 1. Nevertheless, the PACF decays slowly and oscillates at lags $h = 12k$, for $k = 1, 2, \cdots$, and exhibits "tailing off" instead of "cutting off." In this regard, $P$ should be equal to 0.

Now only the value of $q$ is absent in the ARIMA $(2, 0, q) \times (0, 1, 1, 12)$ model. To this end, the cross-validation is designed to find the best ARIMA model that gives the lowest Akaike information criterion (AIC, [29]). The total set of 492 months is split into a training set and a test set in five different ways. The size of the test size is always 60 months, while the training set sizes vary from 60 to 432 months (Figure 12), each of fold uses the same test set size but monotonically increasing training set sizes, the test sets cover the period from month 432 to month 492. The value $q = 3$, which gives the lowest mean AIC among the cross-validations, is finally selected as the best candidate for the ARIMA model.

The model is then fitted with the in-sample data to compute the parameters in Equation (11). The size of the in-sample data set will influence the goodness of fit and

the accuracy of the forecast. To choose a suitable in-sample size (i.e., training set size), we compare the RMSEs between the forecast SICs (obtained from training sets) and the true SICs (in test sets) in Figure 13, in each cross-validation, the model is first fitted by data in the training set. Then, the fitted model forecasts the SICs in the months covered by the test set. The RMSE represents the difference between the forecast and the true SICs in the same period. As the training set size increases, the RMSEs decrease monotonically. For training set sizes larger than 300 months, the RMSE curve shows convergence, following the recommended in-sample size in [30]. Eventually, 432 months are chosen as the in-sample size.
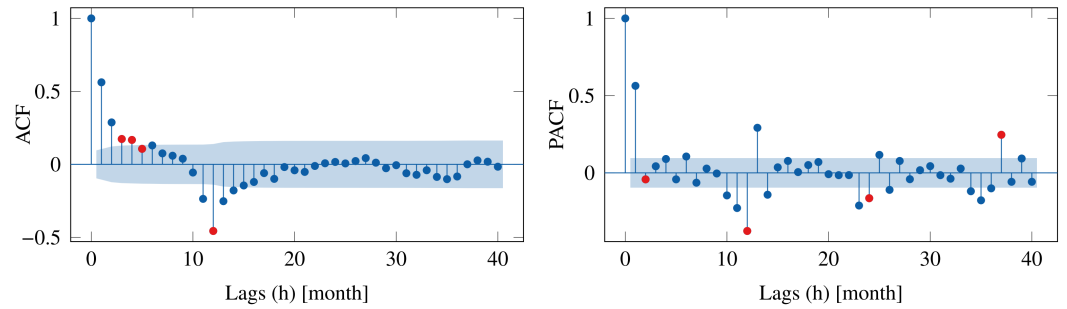


**Figure 11.** The ACF and PACF of the seasonally differenced SIC series in sub-region No. 2. The blue shaded regions represent the range of zero plus/minus two standard errors. When values fall within the shaded region, the ACF and PACF are considered to be zero. The lags of the red dots provide information to determine the orders of the ARIMA model.
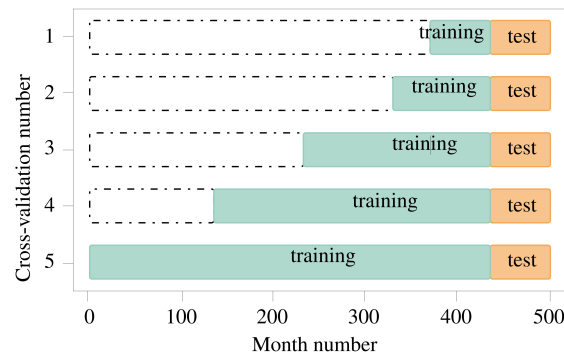


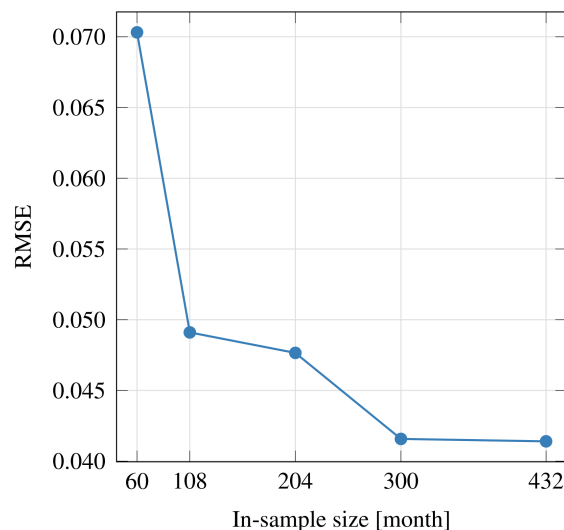**Figure 12.** Schematic of the five folds cross-validations.



**Figure 13.** Convergence study of in-sample sizes.

Using the MLE as described in Section 3.2 to solve for the parameters, the ARIMA $(2, 0, 3) \times (0, 1, 1, 12)$ model is finally expressed as

$$(1 - B^{12})Y_t' =$$
$$\frac{1 - 0.5592B + 0.5290B^2 + 0.3939B^3}{1 - 0.0851B - 0.7502B^2} \left(1 + 0.8520B^{12}\right)\epsilon_t ,$$
$$\epsilon_t \sim N(0, 0.0028).$$

(15)

Recalling the limitations of $\epsilon_t$ in Equation (10), it is important to assess the residual of the ARIMA $(2, 0, 3) \times (0, 1, 1, 12)$ model to ensure its validity. The residual is the difference between the actual SIC values and the values estimated by the model. The quarterly and yearly average residuals are shown in Figure 14, and they oscillate around zero, indicating a white noise pattern. The histogram, Q-Q plot, and autocorrelation function (ACF) of the residuals are plotted in Figure 15. The histogram and Q-Q plot demonstrates that the residuals are almost normal in distribution, and the ACF indicates that the residuals are independent of time, which are the two main assumptions for the residual in the ARIMA model. Hence, the ARIMA $(2, 0, 3) \times (0, 1, 1, 12)$ model has effectively captured most of the information from the original SIC series and the residual is, in essence, white noise.
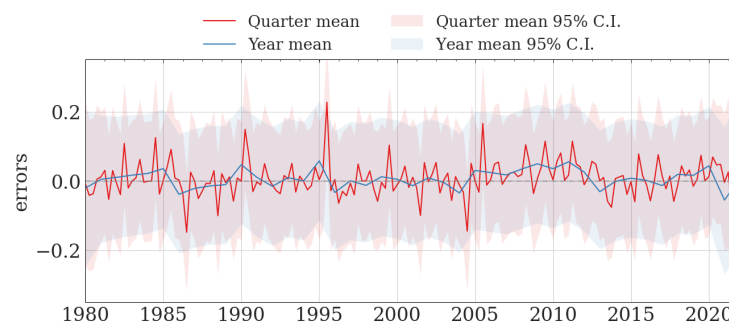


**Figure 14.** The quarterly and yearly mean of the residual.
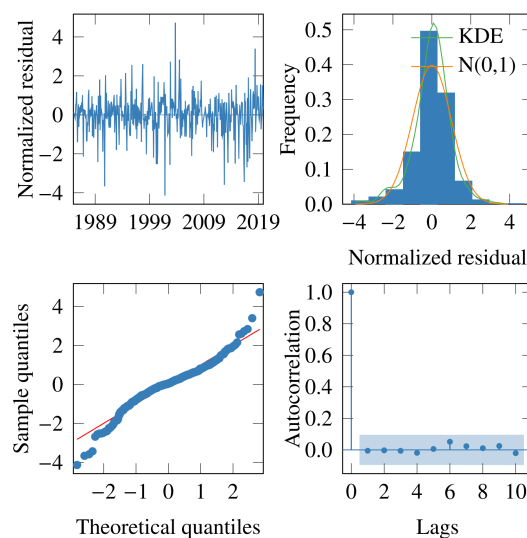


**Figure 15.** The residual diagnosis of the proposed ARIMA model.

Since the ARIMA model has demonstrated acceptable goodness-of-fit, we employ the MMSE forecast [28] to project the SIC from 2021 to 2025 (in Figure 16). The red lines depict the input data obtained from reanalysis, while the blue dotted lines serve as a reference for verification. The green lines represent the SIC values estimated via the ARIMA model, and the purple lines indicate the forecasts made by the same model. The red shaded areas

denote the 95% confidence intervals of the forecasts. Prior to 2021, the estimated SIC values are nearly identical to the input data from reanalysis. Additionally, the forecasted SIC values for the year 2021 align well with the reanalysis data, signifying that the proposed ARIMA model is capable of generating reliable predictions of SIC.
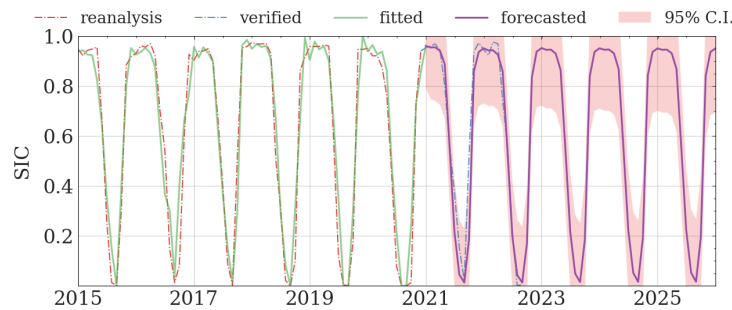


**Figure 16.** The estimated and forecasted SIC values from the ARIMA model in sub-region No. 2.

Furthermore, the quarterly means of the estimated and forecasted SIC values are illustrated in Figure 17. The red lines denote the reanalysis data, while the blue curves represent the SIC values from the ARIMA model. The red shaded areas represent the 95% confidence intervals of forecasts. The plotting of the quarterly means of SIC values allows for easier discernment of the changing trends of the SIC across the quarters. Throughout all quarters, the estimated SIC values align with the reanalysis data, effectively capturing the declining trend of SIC values. The largest discrepancy occurs in the October to December quarter (OND) of the year 2021. Nonetheless, the reanalysis SIC remains within the confines of the 95% confidence intervals.
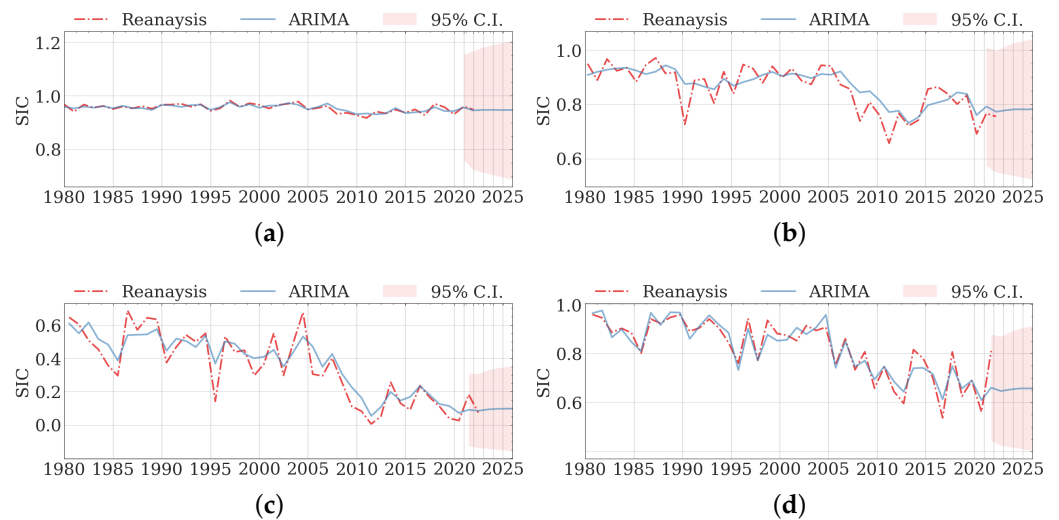


**Figure 17.** The quarterly means of reanalysis and forecast SICs in sub-region No. 2. (**a**) January, February, and March (JFM). (**b**) April, May, and June (AMJ). (**c**) July, August, and September (JAS). (**d**) October, November, and December (OND).

The distributions of quarterly means of SIC values across the various sub-regions are depicted in Figure 18. It is observed that the SICs in sub-region No. 1 are notably lower than in other sub-regions, reaching close to zero from October to December. Conversely, higher SICs are observed in the other sub-regions from January to March and April to June. Additionally, a discernible downward trend in SICs is evident from July to September. The forecasting of the SIC values reveals that the fluctuations are less pronounced than in the reanalysis data, yet the trends are properly captured.
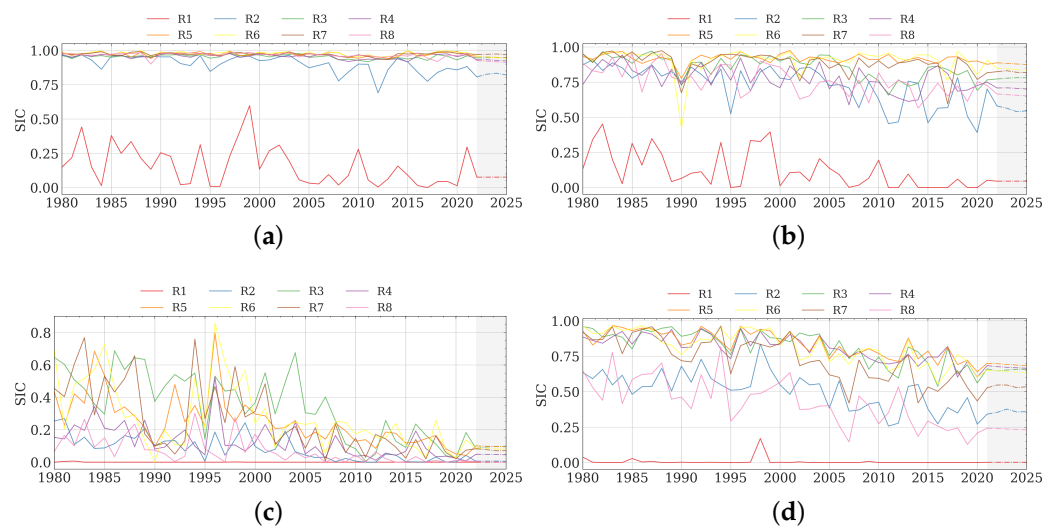
**Figure 18.** The quarterly means of reanalysis and forecast SICs in all sub-regions. (**a**) January, February, and March (JFM). (**b**) April, May, and June (AMJ). (**c**) July, August, and September (JAS). (**d**) October, November, and December (OND).

### 4.3. Validation of SIC Forecasts

We have shown the high goodness-of-fit of the proposed ARIMA model. Nevertheless, the forecast SICs need further validation. To this end, SICs are selected from nine CMIP5 experiments [23]. Although the CMIP5 lacks data before 2006, both CMIP5 and ERA5 provide SICs from January 2006 to September 2022. Therefore, this period is chosen to compare the accuracy of the proposed model. Figure 19 identifies the absolute errors of the monthly averages of SICs between CMIP5 and ERA5. Each dot represents an absolute error in a single month, and the color represents the quarter to which the month belongs. Figure 20 compares the RMSE of the monthly average of SICs between CMIP5 and ERA5. The most accurate experiment is listed at the top and the least accurate experiment at the bottom. The top four experiments are selected as the reference to validate the forecast SICs of the proposed ARIMA model.
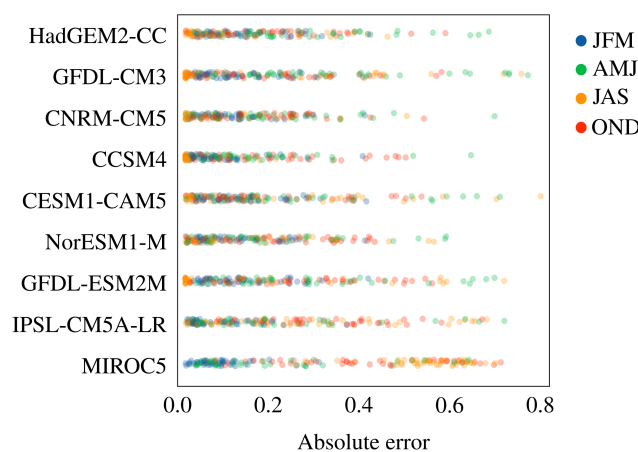


**Figure 19.** The absolute errors of SICs between CMIP5 and ERA5 from January 2006 to September 2020 in sub-region No. 2.
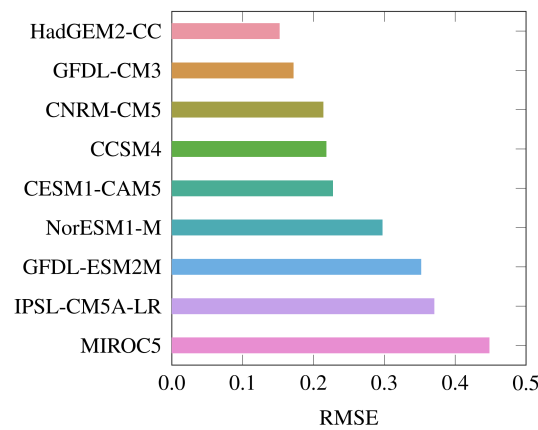
**Figure 20.** The RMSEs of SICs between CMIP5 and ERA5 from January 2006 to September 2020 in sub-region No. 2.

In October, November, and December in sub-region No. 2, the monthly averaged SICs of the selected experiments are plotted in Figure 21. The solid blue line represents the mean SICs of these four experiments, surrounded by the 95% confidence interval (green shaded region). The ERA5 SICs are plotted as the solid brown line, and the ARIMA model regressed SICs are represented as the dashed blue line. From 2006 to 2015, the mean experimental SIC deviates from the SICs of ERA5 and the ARIMA model. Between 2015 and 2022, all three lines lie close together. From 2023, the ARIMA model forecast SICs stay close to the mean SICs of the CMIP5 experiments, the differences of which are bounded by 0.1 of the SIC value. Therefore, we can conclude that the proposed ARIMA model can provide the same confidence for forecasts as CMIP5 experiments.
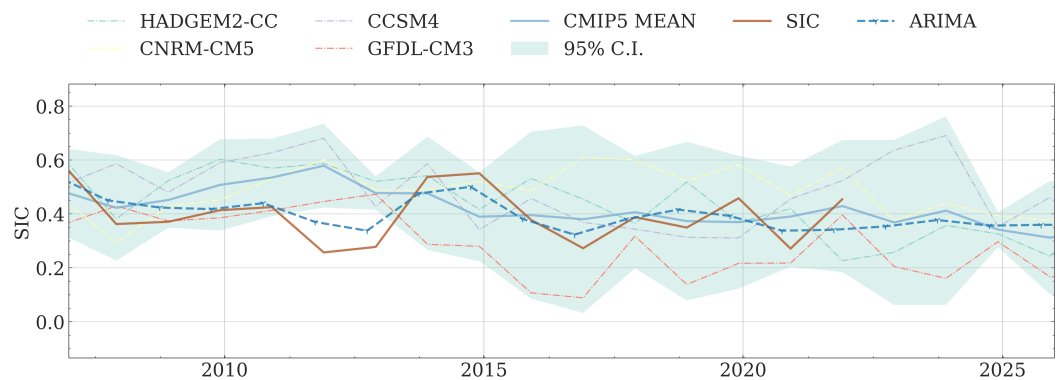


**Figure 21.** October, November, and December mean SICs from ERA5, CMIP5, and the ARIMA model in sub-region no. 2.

### 4.4. Estimation of the Transit Navigation Window

The reanalysis and ARIMA-model-forecasted SICs serve as the basis for determining the navigational window for transit. SIC values that fall below 0.15 or 0.25 are considered to mark the threshold of ice-free conditions, as suggested in [31]. Hence, an 80% navigational window for transit is defined as the number of days per year in which the SIC values along 80% of the route surpass the threshold of ice-free. For any given date, if the SICs in six of the eight sub-regions fall below the threshold, that date is deemed to be included within the navigational window.

Pursuant to this principle, we proceed to estimate the 80% transit navigation window from the year 2022 to 2025. Figure 22 portrays the reanalysis and forecast transit navigation windows in conjunction. The navigation window that results from the use of a higher ice-free threshold is wider than the estimate obtained with a lower threshold, though the trends of both navigation windows are consistent. The navigation windows undergo cyclical

fluctuations from wide to narrow from the year 1991 to 2025. Prior to 2013, the navigation window would experience contraction approximately every seven years, in 1999, 2007, and 2013. Since then, the navigation window has remained periodically vibrational, yet the narrowest windows are now noticeably broader than in the past. In 2021, the navigation window commenced in early August and terminated in mid-September (when using a high SIC threshold), implying a relatively narrow window. By the year 2022, the commencement of the window was postponed by approximately two weeks, but the window was extended until October. Since 2023, the navigation window will broaden, remaining from mid-July to the end of September. The navigation window will continue to expand slightly until 2025.

To validate the estimated navigation window using SICs from the statistical interpolation, we also utilized the SICs from the kriging interpolation to assess the navigation windows. As demonstrated in Figure 23, the navigation windows obtained from both kriging and statistical interpolations are in agreement for the period from 2016 to 2021. This convergence attests to the suitability and veracity of the proposed methodology for estimating the transit navigation window.
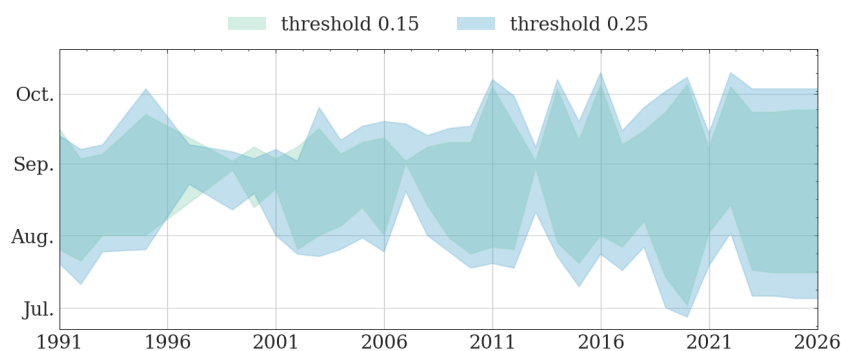


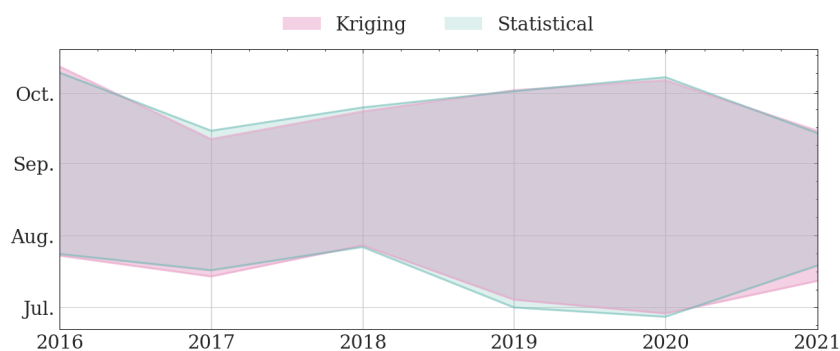**Figure 22.** Estimation of 80% ice-free transit navigation window for the selected NSR.



**Figure 23.** Comprison of 80% ice-free transit navigation window utilizing SICs from statistical and kriging interpolations.

## 5. Conclusions

This study introduces a systematic method for forecasting Sea Ice Concentrations along the Northern Sea Route for Arctic shipping and evaluating the transit navigation window. The Arctic region is divided into sub-regions that encompass the NSR segments, and the mean SIC of each NSR segment in each sub-region is obtained through the proposed statistical interpolation. This technique is found to be less sensitive to grid size variations than traditional nearest grid point interpolation, and yields stable results even with low grid resolution, as it incorporates all statistical information within the sub-region. Moreover, this proposed statistical interpolation has been demonstrated to exhibit comparable accuracy to the widely employed kriging interpolation method.

The interpolated SICs are subsequently analyzed and modeled through the use of an Auto Regressive Integrated Moving Average model, which proves capable of producing

accurate short-term SIC forecasts. This purely statistical model, which only relies on re-analysis (historical) data, has a forecast accuracy comparable to that of the more complex Climate Model Intercomparison Project 5 physical models, which require substantial meteorological and oceanographic data. Moreover, the ARIMA model is relatively simple to understand and implement, and is computationally efficient, making it an ideal candidate for real-time SIC estimation systems in onshore and offshore environments. The ARIMA model can be updated in real-time by re-training it with new SIC data, allowing for dynamic correction of forecasts, and continuous creation of new forecasts as the training set expands.

Finally, the SIC forecasts are applied to evaluate the transit navigation window for the years to come, indicating that the selected NSR is poised to become readily accessible to shipping, with improved sea ice conditions in the years ahead. In this investigation, we selected a grid size of 0.25° by 0.25° in order to assess the transit navigation window and authenticate the proposed statistical models. Although this grid resolution is relatively large and may augment the uncertainty in estimating the transit navigation window, the utilization of high-resolution sea ice information through the proposed statistical technique can yield a more precise transit navigation window. The comparison of uncertainty between high and low resolution represents an intriguing topic and leaves further examination.

The present study proposes a methodology that presents a reliable and efficient approach to forecasting sea ice conditions along the Northern Sea Route. The suggested technique is characterized by its explicitness and low computational requirements, which enable its successful implementation both onboard and onshore. The auto-regression model's inherent characteristics facilitate continual learning and updating, thereby providing a sound basis for weather forecast justifications. The proposed methodology has significant implications for shipping companies in terms of managing and scheduling their Arctic shipping itineraries. By providing a preliminary estimate of the navigation window, the methodology allows companies to optimize their shipping operations along the NSR, resulting in reduced risks of delay and financial loss.

**Author Contributions:** Conceptualization, D.W., W.T., X.L. and W.M.; methodology, D.W.; software, D.W.; validation, D.W., X.L., W.T. and W.M.; formal analysis, D.W.; investigation, D.W.; resources, D.W. and J.Z.; data curation, D.W.; writing—original draft preparation, D.W.; writing—review and editing, X.L., W.T., W.M. and J.Z.; visualization, D.W.; supervision, W.M.; project administration, J.Z.; funding acquisition, D.W. and J.Z. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

# References

1. Milaković, A.S.; Gunnarsson, B.; Balmasov, S.; Hong, S.; Kim, K.; Schütz, P.; Ehlers, S. Current Status and Future Operational Models for Transit Shipping along the Northern Sea Route. *Mar. Policy* **2018**, *94*, 53–60. [CrossRef]
2. Zhang, C.; Zhang, D.; Zhang, M.; Mao, W. Data-Driven Ship Energy Efficiency Analysis and Optimization Model for Route Planning in Ice-Covered Arctic Waters. *Ocean Eng.* **2019**, *186*, 106071. [CrossRef]
3. Lang, X.; Zhang, C.; Jonasson, L.; Mao, W.; Eriksson, L.; Zhang, D. Comparison between Full-Scale Measurements and Theoretical Fuel Consumption Model in a Real Arctic Ship Navigation. In Proceedings of the 29th International Ocean and Polar Engineering Conference, Honolulu , HI, USA, 15–20 June 2019.
4. Zhang, C.; Zhang, D.; Zhang, M.; Lang, X.; Mao, W. An Integrated Risk Assessment Model for Safe Arctic Navigation. *Transp. Res. Part A Policy Pract.* **2020**, *142*, 101–114. [CrossRef]

5. Chen, J.; Kang, S.; Du, W.; Guo, J.; Xu, M.; Zhang, Y.; Zhong, X.; Zhang, W.; Chen, J. Perspectives on Future Sea Ice and Navigability in the Arctic. *Cryosphere* **2021**, *15*, 5473–5482. [CrossRef]

6. Liu, X.H.; Ma, L.; Wang, J.Y.; Wang, Y.; Wang, L.N. Navigable Windows of the Northwest Passage. *Polar Sci.* **2017**, *13*, 91–99. [CrossRef]

7. Zhou, X.; Min, C.; Yang, Y.; Landy, J.C.; Mu, L.; Yang, Q. Revisiting Trans-Arctic Maritime Navigability in 2011–2016 from the Perspective of Sea Ice Thickness. *Remote Sens.* **2021**, *13*, 2766. [CrossRef]

8. Inoue, J. Review of Forecast Skills for Weather and Sea Ice in Supporting Arctic Navigation. *Polar Sci.* **2021**, *27*, 100523. [CrossRef]

9. Li, Z.; Ryan, C.; Huang, L.; Ding, L.; Ringsberg, J.W.; Thomas, G. A Comparison of Two Ship Performance Models against Full-Scale Measurements on a Cargo Ship on the Northern Sea Route. *Ships Offshore Struct.* **2021**, *16*, 237–244. [CrossRef]

10. Stark, J.D.; Ridley, J.; Martin, M.; Hines, A. Sea Ice Concentration and Motion Assimilation in a Sea Ice-Ocean Model. *J. Geophys. Res. Ocean.* **2008**, *113*, C05S91. [CrossRef]

11. Ye, Y.; Shokr, M.; Heygster, G.; Spreen, G. Improving Multiyear Sea Ice Concentration Estimates with Sea Ice Drift. *Remote Sens.* **2016**, *8*, 397. [CrossRef]

12. Yu, L.; Zhong, S.; Winkler, J.A.; Zhou, M.; Lenschow, D.H.; Li, B.; Wang, X.; Yang, Q. Possible Connections of the Opposite Trends in Arctic and Antarctic Sea-Ice Cover. *Sci. Rep.* **2017**, *7*, 45804. [CrossRef] [PubMed]

13. Ye, Y.; Shokr, M.; Aaboe, S.; Aldenhoff, W.; Eriksson, L.E.B.; Heygster, G.; Melsheimer, C.; Girard-Ardhuin, F. Inter-Comparison and Evaluation of Sea Ice Type Concentration Algorithms. *Cryosphere Discuss.* **2019**, 1–30. [CrossRef]

14. Mu, L.; Liang, X.; Yang, Q.; Liu, J.; Zheng, F. Arctic Ice Ocean Prediction System: Evaluating Sea-Ice Forecasts during Xuelong's First Trans-Arctic Passage in Summer 2017. *J. Glaciol.* **2019**, *65*, 813–821. [CrossRef]

15. Spanos, P.T.D. ARMA Algorithms for Ocean Wave Modeling. *J. Energy Resour. Technol.* **1983**, *105*, 300–309. [CrossRef]

16. Soares, C.G.; Ferreira, A.M.; Cunha, C. Linear Models of the Time Series of Significant Wave Height on the Southwest Coast of Portugal. *Coast. Eng.* **1996**, *29*, 149–167. [CrossRef]

17. Agrawal, J.D.; Deo, M.C. On-Line Wave Prediction. *Mar. Struct.* **2002**, *15*, 57–74. [CrossRef]

18. Jiang, H.; Duan, S.; Huang, L.; Han, Y.; Yang, H.; Ma, Q. Scale Effects in AR Model Real-Time Ship Motion Prediction. *Ocean Eng.* **2020**, *203*, 107202. [CrossRef]

19. Mao, W.; Rychlik, I.; Wallin, J.; Storhaug, G. Statistical Models for the Speed Prediction of a Container Ship. *Ocean Eng.* **2016**, *126*, 152–162. [CrossRef]

20. Piwowar, J.M.; LeDrew, E.F. On the Auto Regressive Nature of Arctic Sea Ice Concentrations. *Can. J. Remote Sens.* **2001**, *27*, 517–525. [CrossRef]

21. Ahn, J.; Hong, S.; Cho, J.; Lee, Y.W.; Lee, H. Statistical Modeling of Sea Ice Concentration Using Satellite Imagery and Climate Reanalysis Data in the Barents and Kara Seas, 1979–2012. *Remote Sens.* **2014**, *6*, 5520–5540. [CrossRef]

22. Copernicus Climate Change Service Climate Data Store (CDS). ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. 2023. Available online: https://cds.climate.copernicus.eu (accessed on 6 February 2023).

23. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An Overview of CMIP5 and the Experiment Design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [CrossRef]

24. Baumohl, E.; Lyocsa, S. *Stationarity of Time Series and the Problem of Spurious Regression*; SSRN Scholarly Paper ID 1480682; Social Science Research Network: Rochester, NY, USA, 2009.

25. Mushtaq, R. *Augmented Dickey Fuller Test*; SSRN Scholarly Paper ID 1911068; Social Science Research Network: Rochester, NY, USA, 2011.

26. Shiskin, J. *The X-11 Variant of the Census Method II Seasonal Adjustment Program*; Number 15; US Department of Commerce, Bureau of the Census: Washington, DC, USA, 1967.

27. Cordeiro, G.M.; Klein, R. Bias Correction in ARMA Models. *Stat. Probab. Lett.* **1994**, *19*, 169–176. [CrossRef]

28. Box, G.E.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis, Forecasting and Control*; Prentice Hall: Englewood Cliffs, NJ, USA, 1994.

29. Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]

30. Shumway, R.H.; Stoffer, D.S.; Stoffer, D.S. *Time Series Analysis and Its Applications*; Springer: Cham, Switzerland, 2000; Volume 3.

31. 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*, 024010. [CrossRef]