Journal of *Marine Science and Engineering*

MDPI

Article

Identifying Key Spatiotemporal Regions of the Local Source of the Northern Yellow Sea Cold Water Mass

Xiao Chen [1,2,3,*], Zuozuo Ma [1,2], Miangang Song [1,2], Zhiliang Liu [1,2], Tao Liu [1,4], Yunlong Lu [1,2] and Jia Shi [1,4]

[1] Research Center for Marine Science, Hebei Normal University of Science and Technology, Qinhuangdao 066004, China
[2] Hebei Key Laboratory of Ocean Dynamics, Resources and Environments, Qinhuangdao 066004, China
[3] Qinhuangdao Key Laboratory of Marine Ecological Security and Artificial Intelligence Application, Qinhuangdao 066004, China
[4] School of Mathematics and Information Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao 066004, China
* Correspondence: chenxiao0604@hevttc.edu.cn

Abstract

The Northern Yellow Sea Cold Water Mass (NYSCWM) is a distinctive hydrographic phenomenon in China's coastal waters and is generally considered to originate from locally formed cold water during the previous winter. However, the specific wintertime period and local spatial range controlling its bottom-layer minimum temperature (BMT) remain unclear. This study utilizes August BMT data spanning 2003–2020, together with winter Multiscale Ultrahigh Resolution Sea Surface Temperature (MURSST) data. On this basis, K-means clustering is applied to identify the key spatiotemporal regions linked to BMT variability. Results show that the BMT of the NYSCWM exhibits a significant warming trend of about 0.0533 °C yr^{-1} and a pronounced quasi-3-year oscillation. The strongest correlation (CC = 0.8396) between BMT and winter SST occurs in the central Northern Yellow Sea (NYS) during the second half of February, exceeding that in other regions. This area acts as a key spatiotemporal region, located between colder western waters and warmer southern sectors, and maintains persistently low temperatures during this period. A regression model based on SST in this key spatiotemporal region reproduces observed BMT with a correlation coefficient of 0.9146 and enables prediction six months in advance. These results refine the identification of key spatiotemporal regions and improve our understanding of NYSCWM formation and evolution.

Keywords: Northern Yellow Sea Cold Water Mass; bottom-layer minimum temperature; local source; winter SST; key spatiotemporal regions; K-means clustering; regression model

Academic Editor: Jean-Louis Pinault

Received: 18 April 2026
Revised: 12 May 2026
Accepted: 13 May 2026
Published: 15 May 2026

1. Introduction

The Yellow Sea Cold Water Mass (YSCWM) is a distinctive hydrological phenomenon occurring in the shallow shelf seas of China [1,2]. It forms through the strong cooling and vertical mixing of local water in winter. With the strengthening of thermal stratification, the YSCWM persists in the deep bottom layer of the central Yellow Sea during summer [3–5]. The YSCWM is a distinct water mass characterized by low temperature and high salinity [6], covering an area of approximately 130,000 km^2 and occupying a volume of about 500 billion m^3. This region is nutrient-rich and exhibits high water quality, supporting abundant meiobenthos and providing an ideal habitat for salmonids and other fish species [7,8]. In addition, it underpins a large-scale offshore cold-water fish farming

industry worth hundreds of billions of CNY, and plays a vital role in promoting the marine economy and contributing to the national "Blue Granary".

In summer, the YSCWM typically exhibits two to three distinct cold centers [9–11]. In years with weak cold water mass intensity, such as 1935 and 1972, only one cold center appears in the Northern Yellow Sea (NYS), with the bottom-layer minimum temperature (BMT) generally ranging from approximately 4 to 8 °C [9,12]. Compared with the Southern Yellow Sea (SYS), the Northern Yellow Sea Cold Water Mass (NYSCWM) is relatively stable in position and shows smaller interannual variability. Constrained by topography, its cold center is mainly located in the deepest area of the trough and slightly shifts toward Dalian [10,13]. However, affected by typhoons and other extreme weather events [14,15], the position or boundary of the cold water mass can change sharply, intensifying fluctuations in environmental factors such as nutrient concentration in the adjacent sea area. This hinders the growth of organisms such as Yesso scallops (*Patinopecten yessoensis*) and even causes mortality [16]. Meanwhile, under a sustained warming trend, the cold water mass is gradually weakening, leading to bottom water hypoxia and massive mortality of sea cucumbers and other bottom-seeded organisms [17,18]. These changes in the cold water mass have exerted profound impacts on the regional marine economy, ecosystem structure, and fishery resource distribution. Therefore, it is essential to conduct an in-depth study on the evolutionary characteristics and formation mechanisms of the NYSCWM.

Existing studies indicate that the YSCWM exhibits significant interannual and interdecadal variability with periods of 2–7 years and 10–20 years [11,19,20], respectively. It is primarily influenced by global-scale climate modes, including El Niño–Southern Oscillation (ENSO), Arctic Oscillation (AO), and Pacific Decadal Oscillation (PDO) [21–24]. Among them, the interannual variability of the NYSCWM is closely related to the coupled interaction between ENSO and atmospheric circulation, exhibiting a 3–5 year oscillation cycle. Data from the Dalian–Chengshantou (abbreviated as DC) transect during 1976–2019 indicate that the BMT of the NYSCWM shows an overall increasing trend, with a warming rate ranging from about 0.005 to 0.026 °C yr^{-1} [21,23–25]. The minimum value occurred in 2011, at approximately 4.93 °C. Subsequently, a significant low-frequency warming trend has led to the continuous weakening of the NYSCWM in recent years.

Furthermore, the NYSCWM also shows a significant response to local climate conditions and air–sea interaction processes during the previous winter. Early observational data indicate that the correlation coefficient is approximately 0.62 between the BMT of the DC transect in August and the local mean air temperature in January. The air temperature exhibits a significant quasi-biennial (2–3 years) oscillation, and the BMT shows a corresponding oscillatory variability [26–28]. Since the 21st century, studies have been conducted based on extensive marine observational data and numerical simulations. These studies have found that changes in the mean water temperature, BMT, and boundary of the NYSCWM are closely related to local factors, such as preceding winter Sea Surface Temperature (SST), air temperature, winter monsoon, surface radiative flux over the study region, and the influence of the Yellow Sea Warm Current (YSWC) [19,21,25,29,30]. This further indicates that the intensity of the summer cold water mass primarily depends on the gain or loss of heat content in the region during the previous winter [21,31]. This is consistent with the formation mechanism of the cold water mass driven by intense surface cooling and deep vertical mixing during winter [14,32].

Overall, existing studies have comprehensively described the formation and interannual variability characteristics of the NYSCWM, as well as its large-scale and local air–sea influencing factors. They have provided important references for in-depth understanding of the evolutionary mechanism of the cold water mass. These studies further support the widely accepted view that the NYSCWM originates from locally formed cold water during

the previous winter, reflecting a typical lagged response from winter to summer. Therefore, identifying and understanding the variability of previous winter SST in key local regions is crucial for explaining and predicting changes in NYSCWM intensity. However, the existing studies still have several limitations. Most analyses are based on low-resolution, short-term, and discontinuous observations; most focus on the correlation analysis between NYSCWM and regional overall winter SST, without identifying the key regions of the NYSCWM's local source. They also fail to explore the potential influence of wintertime synoptic-scale variability on cold water mass.

In summary, based on long-term continuous NYSCWM observations and ultra-high-resolution MURSST data from 2003 to 2020, this study systematically investigated the correlation between summer BMT along the DC transect and the previous winter SST. The aim was to accurately identify the local core regions and critical time windows dominating BMT variations using K-means clustering. Subsequently, a regression model was constructed to fit and predict BMT, thereby verifying the robustness of the identified spatiotemporal regions. The results of this study can not only deepen the understanding of the interannual variability of the NYSCWM and its response to previous winter air–sea processes, but also provide methodological guidance for the refined identification of key sources and causes. This further provides important support for early warning of abnormal changes in cold water mass, rational management of fishery resources, and scientific prevention and control of ecological and environmental risks.

2. Data and Methods

2.1. Data

2.1.1. Research Area and Data

The research area is the NYS, primarily located between 37° N and 40° N and 121° E–126° E, as shown in Figure 1. The red dotted line marks the DC transect, which is the main observation section of the NYSCWM; the low-temperature center (hereafter referred to as the cold center) of the NYSCWM in August is primarily located near this transect.

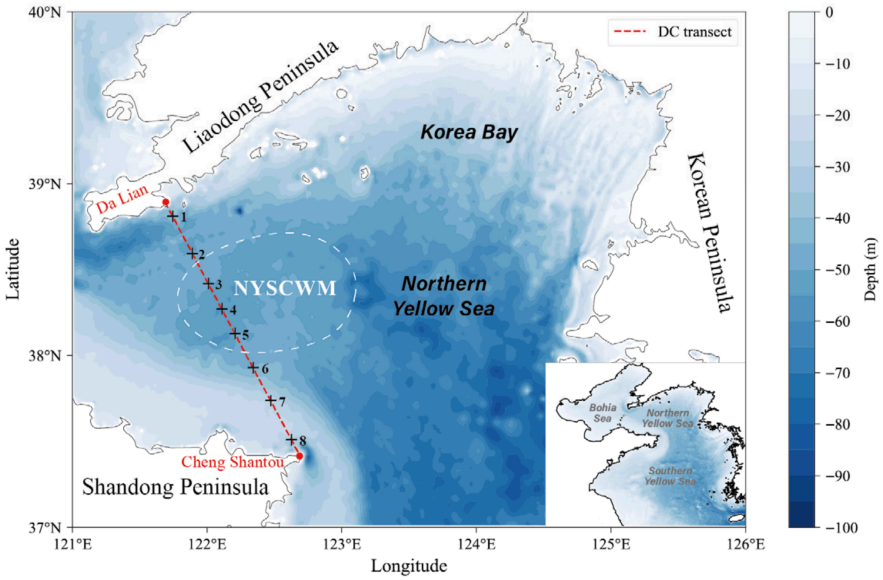

Figure 1. Research area of the Northern Yellow Sea and the DC transect lines. The red dashed line represents the DC transect, which consists of 8 sampling stations (1–8) deployed from Dalian (northern end) to Chengshantou (southern end). The arrow indicates the transect direction. The white dashed circle marks the typical distribution range of the Northern Yellow Sea Cold Water Mass (NYSCWM). The background shows the bathymetric map of the Northern Yellow Sea.

This study mainly focuses on the BMT data of the cold center in the NYS during 2003–2020. The data are derived from two primary sources, covering 2003–2010 and 2011–2020, respectively: First, the 2003–2010 data were compiled from published historical studies, including in situ observations of the YSCWM collected over several decades. Second, the 2011–2020 data were provided by the North China Sea Environmental Monitoring Center, State Oceanic Administration, and were obtained during annual summer comprehensive oceanographic surveys using shipborne Conductivity–Temperature–Depth (CTD) profilers. The CTD observations included vertical temperature and salinity profiles from the surface to near-bottom layers. The locations of the CTD stations are indicated by black circles in Figure 1.

2.1.2. Northern Yellow Sea SST Data

The Multiscale Ultrahigh Resolution Sea Surface Temperature (MURSST) dataset originates from the Jet Propulsion Laboratory (JPL) of the National Oceanic and Atmospheric Administration (NOAA). MURSST is a data product generated through optimal interpolation based on multiple satellite observations, including MODIS, AVHRR, AMSR-E, and WindSat, among others. The literature [33] provides a detailed account of the dataset's sources and fusion methods. MURSST provides global ocean coverage with a spatial resolution of $0.01^\circ \times 0.01^\circ$ and a daily temporal resolution, spanning from June 2002 to the present. This study selected data from the NYS region spanning December 2002 to February 2025 to investigate the interannual variability of SST and the influence of previous winter SST variations on the intensity of the cold center. Previous studies have demonstrated that MURSST performs well in the Northern Yellow Sea, providing reliable SST estimates and serving as a robust data source for long-term prediction and analysis [34].

2.2. Methods

2.2.1. K-Means Clustering Analysis

The K-means clustering [35] is a widely used unsupervised learning method for identifying distinct water masses and analyzing their spatiotemporal characteristics [36,37], thereby revealing their spatial distribution and temporal evolution. The algorithm partitions a dataset into K non-overlapping clusters, ensuring that data points within each cluster are closely grouped while clusters remain well separated. Through iterative optimization, the algorithm minimizes the sum of squared distances (Euclidean distance) between each data point and the centroid of its assigned cluster, achieving automatic data clustering.

Given a dataset $X = \{x_1, x_2, \ldots, x_n\}$ (where n denotes the sample size) and a predefined number of clusters K, the clustering objective function J is defined as shown in Equation (1):

$$J_{Min} = \sum_{k=1}^{K} \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (1)$$

In Equation (1), C_k denotes the set of data points in the $k-th$ cluster, μ_k represents the centroid of the $k-th$ cluster, and $\|x_i - \mu_k\|^2$ corresponds to the squared Euclidean distance. The algorithm iteratively updates cluster assignments and centroid positions, minimizing J until it converges to a local minimum. Previous studies have largely focused on water mass analysis using T–S data or single SST datasets, without considering spatial factors, which has limited understanding of their structural characteristics. To address this, this study integrates SST, geographic coordinates, and seafloor topography and applies K-means clustering on February winter data from 2003 to 2020 on two temporal scales. First, all winter daily data are clustered collectively to characterize the overall spatial pattern of the cold center. Second, multi-year daily data are clustered separately to capture the dynamic temporal evolution of the cold center within the winter season. By comparing these two

scales, the study identifies key spatiotemporal regions that influence the formation and stability of the cold center.

2.2.2. Regression Analysis

Regression modeling [38] is a classical statistical approach that has been widely applied in climate, oceanographic, and environmental studies [39,40]. Compared with linear regression, quadratic polynomial regression not only captures nonlinear relationships between independent and dependent variables but also accounts for potential interactions among the predictors. Therefore, this method is particularly appropriate for quantitatively evaluating the nonlinear and interaction effects of February SST in the NYS on the BMT of the NYSCWM.

A quadratic polynomial regression model, as expressed in Equation (2):

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j + \varepsilon \quad (2)$$

in Equation (2), y denotes serves as the dependent variable representing the BMT of the cold center. x_i denotes the independent variables, where $i = 1$ corresponds to a single variable and $i > 1$ corresponds to multiple variables; the x_i can be defined as the SST values of different regions or dates in the NYS during February. The term β_0 is the intercept, the coefficients $\beta_i (i = 1, 2, \ldots, n)$ represent the linear terms, the coefficients $\beta_{ii} (i = 1, 2, \ldots, n)$ correspond to the quadratic terms, and $\beta_{ij} (1 \leq i < j \leq n)$ represent the coefficients of the interaction terms between independent variables. The error term ε accounts for the unexplained variance in the model. All coefficients are estimated using the least squares method, which enables a quantitative assessment of the influence of February SST on the intensity of the NYSCWM and provides a comprehensive explanation of the interannual variability of the BMT.

3. Results

3.1. Variabilities of the Summer NYSCWM

The cold center of the NYSCWM is primarily located near the DC transect. The BMT serves as a key indicator of cold water mass intensity and effectively reflects the summer variability of the NYSCWM. Therefore, this study focuses on analyzing the BMT along the DC transect as a representative feature of the cold center of the NYSCWM.

The time series of the BMT in summer (August) from 2003 to 2020 is shown in Figure 2a. During these 18 years, the BMT ranged approximately from 4.93 to 8.42 °C, with a mean value of about 5.98 °C. The overall trend exhibits a significant warming trend, and the total warming amplitude was about 0.0533 °C yr^{-1}; notably, the warming rate reached 0.2862 °C yr^{-1} during 2011–2020. This trend is consistent with the results of existing studies [21,25], although the warming rate is stronger than 0.026 °C yr^{-1} during 1976–2006 [21].

According to previous studies, the BMT of the cold center exhibits interannual fluctuations with periods of approximately 2–7 years [20]. Meanwhile, wavelet power spectrum analysis results indicate that the strongest periodic signal of the BMT over the 18-year period is approximately 3 years, as shown in Figure 2b,c. Specifically, during 2003–2010, the variability was dominated by high-frequency oscillations (2–4 years), whereas during 2011–2020, lower-frequency variability (6–8 years) became more prominent, although statistically insignificant. These results indicate that in the most recent decade, the BMT has shifted toward a low-frequency warming regime, suggesting that the intensity of the cold water mass has gradually weakened. This result is generally consistent with previous stud-

ies indicating that the Yellow Sea Cold Water Mass has gradually warmed and weakened under a changing climate [23].

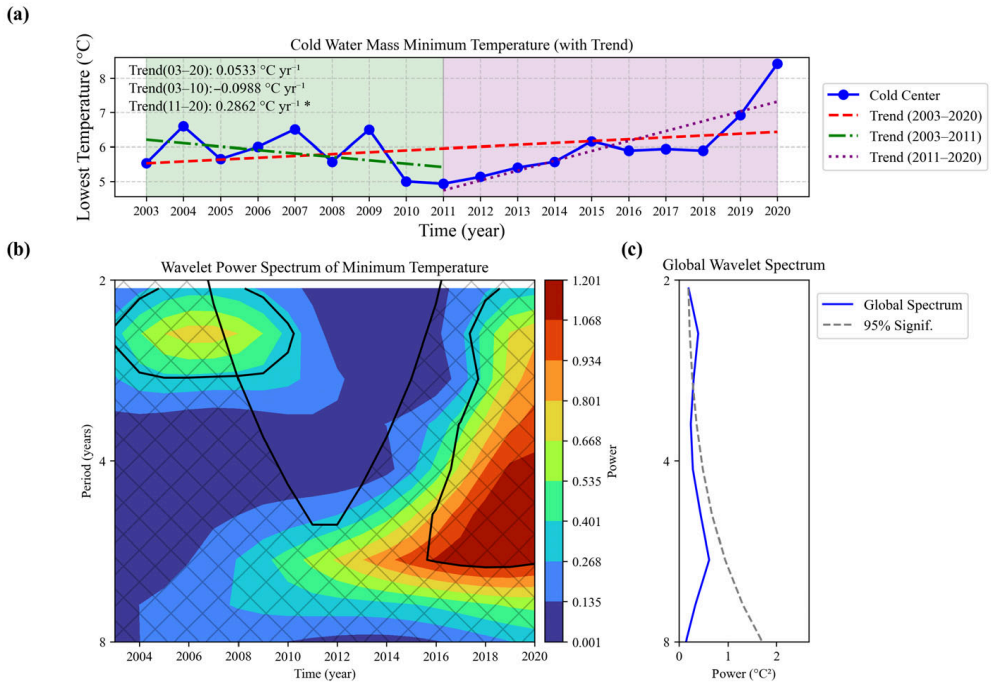

Figure 2. Interannual variation and periodic variability of the cold-center BMT. (**a**) Time series of the BMT at the cold center (* indicates indicates statistical significance at the 95% confidence level); (**b**) wavelet power spectrum, with the thick black contour representing the 95% confidence level; (**c**) global wavelet spectrum, obtained by averaging the wavelet power over the entire time domain, representing the variance distribution across different periodic scales and used to identify the dominant periodic signals in the interannual variability of the cold-center BMT; the blue solid line indicates the actual power spectrum and the orange dashed line represents the significance test against red noise.

3.2. Correlation Analysis Between the BMT and Winter SST

Previous studies have demonstrated that the NYSCWM originates from residual cold water formed through winter cooling and vertical mixing processes [1,26]. To further verify and quantitatively assess the relationship between the two factors, we analyzed the BMT time series at the cold center and the mean winter SST (from December to February) in the NYS during the period 2003–2020, as shown in Figure 3a. The corresponding Pearson correlation coefficients are summarized in Table 1. We can see from the figure and the table that the BMT exhibits a pronounced and statistically significant positive correlation with the SSTs in January, February, and the winter mean, which is consistent with previous studies [21,29]. Among these, the February SST shows the strongest correlation (CC = 0.6963, p = 0.0013), notably exceeding the correlation between the cold center temperature and the local winter air temperature reported in earlier studies [25,26]. Moreover, the February SST displays a distinct interannual periodicity of approximately three years, as shown in Figure 3b, which closely corresponds to the variation cycle of the cold center. Notably, the minimum SST each winter consistently occurs in February, and only in this month does the SST fall entirely below the BMT. These findings collectively suggest that the February SST in the NYS serves as the primary thermal precursor of the cold center BMT, thereby playing a crucial role in determining the intensity and persistence of the YSCWM.

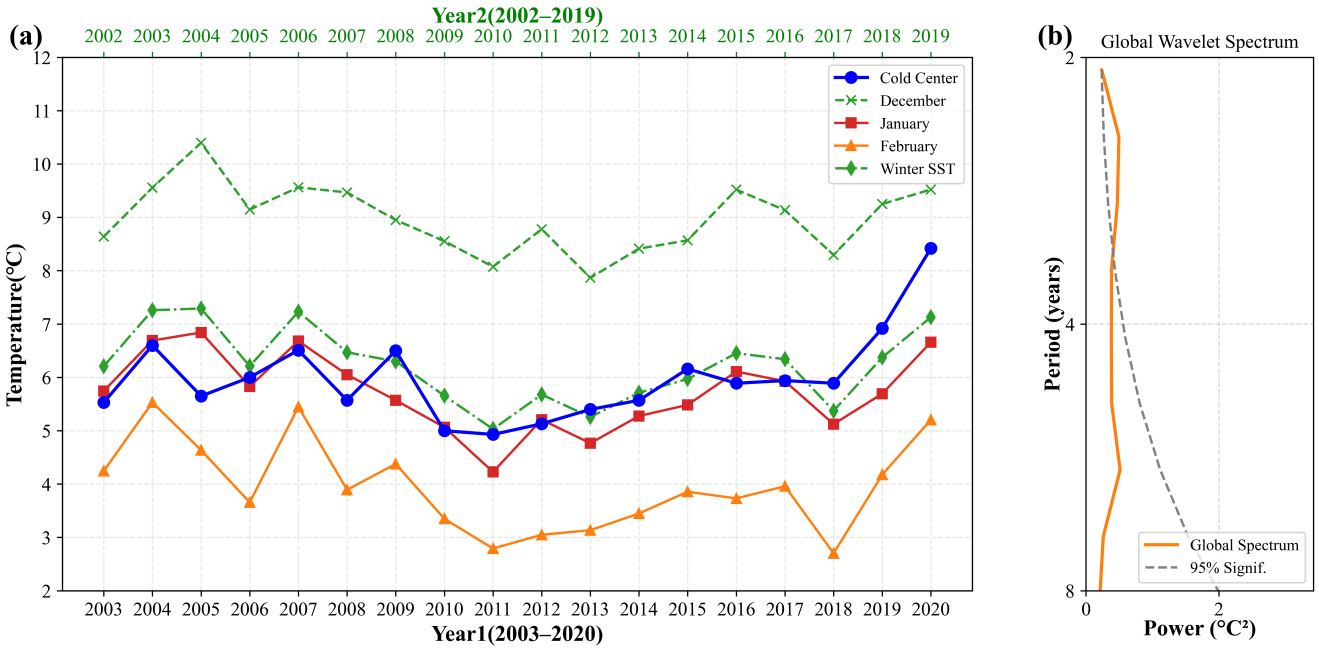

Figure 3. Interannual variations and periodic variability of winter SST. (**a**) Time series of the BMT at the cold center and mean SSTs for December, January, February, and the winter mean; (**b**) global wavelet spectrum of February SST.

Table 1. Correlation between the cold center and the monthly/mean winter SST.

Cold Center (2003–2020)	Correlation Coefficient (CC)	*p*-Value
December (2002–2019)	0.4522	0.0569 (insignificant)
January (2003–2020)	0.6081	0.0074 (significant)
February (2003–2020)	0.6963	0.0013 (significant)
Winter SST (2002–2019)	0.6314	0.0049 (significant)

The spatiotemporal distribution of February SST in the NYS over the 18-year period is shown in Figure 4. The SST ranged from −1.49 °C (2017) to 8.41 °C (2020), exhibiting a pronounced meridional (north–south) distribution pattern, with a northward-extending warm water tongue structure. For example, taking the 5 °C isotherm as an example, its maximum extent reached 39.25° N, 121.00° E (2007), covering an area of approximately 4.1×10^4 km^2. The minimum SST was observed in the northern part of the NYS, ranging from approximately −2 °C to 2 °C, while the maximum SST in the southern region ranged from approximately 6 °C to 9 °C, with the largest temperature difference reaching 9.50 °C (2017). This distribution pattern is primarily controlled by the combined effects of winter monsoon cooling, differential solar radiation, seabed topography, and ocean current structure.

Furthermore, as shown in Figures 3 and 4, during years with relatively high February SST (such as in 2004, 2007, and 2020), that is, when SST $\geq$ 5 °C, the corresponding BMT were also relatively high. Conversely, during years with relatively low February SST (such as in 2011 and 2018), that is, when SST $\leq$ 3 °C, the BMT were lower. Notably, the cold water mass reached its lowest temperature in 2011, which also coincided with a relatively low February SST.

However, for the area-mean SST of the entire domain in February, the maximum occurred in 2004, whereas the highest BMT within the cold center was recorded in 2020. Similarly, the minimum area-mean SST was observed in 2018, yet the BMT did not show a corresponding decrease. This indicates that variations in the area-mean SST of the

NYS cannot fully reflect the variability in the intensity of the cold center; in other words, the entire region does not uniformly control the changes in the BMT. Therefore, it is necessary to identify which specific subregion exerts the dominant control on the cold center's variability.

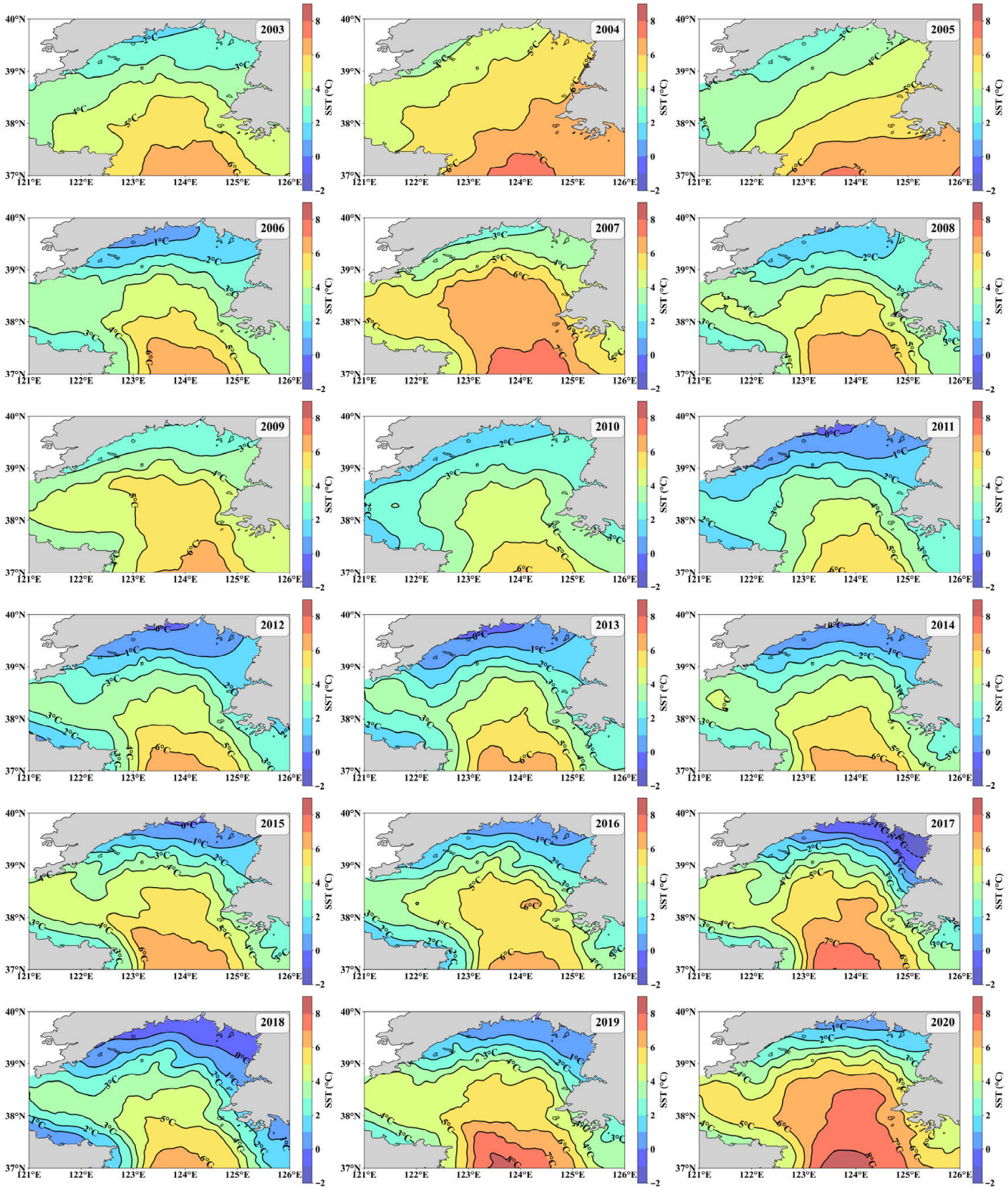

Figure 4. Spatial distribution of February SST for each year.

3.3. Identification of Key Regions

This study comprehensively considers the influence of the February SST, water depth (topography), and latitude–longitude factors on summer cold water masses. The K-means clustering method was applied to the spatial structure of the NYS to identify key regions that govern the evolution of the cold water mass.

This study comprehensively considers the influence of February SST, water depth (topography), and latitude–longitude factors on the summer cold water mass. The K-means clustering method was applied to the spatial structure of the Northern Yellow Sea (NYS) to identify key regions governing the evolution of the cold water mass. To determine the optimal number of clusters, the Elbow Method was employed. The results show a clear inflection point at K = 5, indicating that five clusters represent the optimal classification scheme for the spatial division of the NYS. In addition, the selection of five clusters was not based solely on mathematical results, but also considered the natural geographical boundaries and water mass characteristics of the Northern Yellow Sea to ensure that the clustering results possess clear oceanographic physical significance. Based on February data from 2003 to 2020, the NYS was divided into five regions, as shown in Figure 5a: northern, eastern, southern, western, and central. The northern region, encompassing the southeastern coast of the Liaodong Peninsula and the northwestern coast of the Korean Peninsula (Korea Bay), is strongly influenced by cold air, exhibiting the lowest SST, with a long-term mean of approximately 1.99 °C. The eastern region, located near the western coast of the Korean Peninsula, has slightly higher SST than the northern region. The southern region, situated in the central Yellow Sea, shows the highest SST, with a mean of around 6.04 °C. The western region, covering the northern part of the Shandong Peninsula and parts of the southern coast of the Liaodong Peninsula, corresponds to the DC section and the location of the summer cold center. The central region, located in the central NYS, has a mean SST of about 4.83 °C.

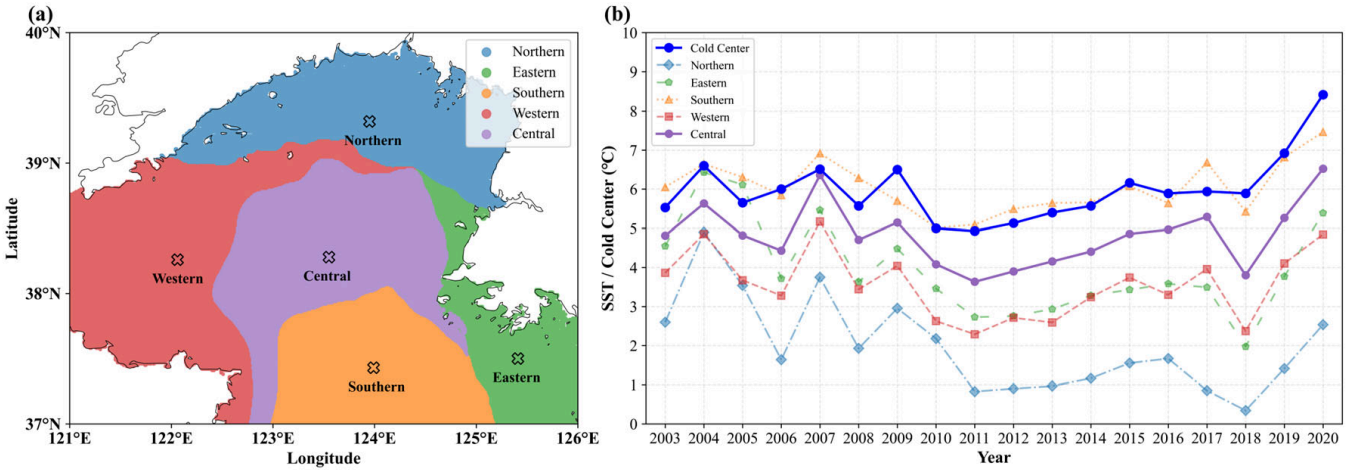

Figure 5. February clustering results in the NYS. (**a**) spatial clusters; (**b**) time series of the cold center and mean SST in five regions. Cluster centroids are denoted by 'X' markers representing the geometric mean position of each cluster.

Based on Figure 5 and Table 2, during the 18-year period, the February SST in the northern, eastern, and western regions exhibited a cooling trend, with cooling rates ranging from −0.0232 to −0.1418 °C yr^{-1}. In contrast, the southern and central regions showed a warming trend, with warming rates ranging from approximately 0.0052 to 0.0140 °C yr^{-1}, which are weaker than that of the summer cold center (0.0533 °C yr^{-1}). Correlation analysis indicates that the cold center has a relatively low correlation with the northern and eastern regions, likely due to the overall significant cooling trend in these regions as revealed

by the Mann–Kendall test. Conversely, the cold center is positively correlated with the western, southern, and central regions (significance tested via *p*-value), with correlation coefficients substantially higher than that for the entire February domain (0.6963). These findings suggest that the southern, western, and central regions form a dynamic system in winter, collectively influencing the intensity of the cold water mass.

Table 2. February SST statistics in five regions and their correlation with the cold center.

Region	Mean SST (°C)	Trend (°C yr^{-1})	Correlation Coefficient (CC)/*p*-Value
Northern	1.9869	−0.1418	0.3725/0.1279 (insignificant)
Eastern	3.9581	−0.1113	0.5195/0.0271 (significant)
Southern	6.0444	0.0140	0.8121/0.0000 (significant)
Western	3.5642	−0.0232	0.7646/0.0002 (significant)
Central	4.8250	0.0052	0.8356/0.0000 (significant)
Entire Region	3.9552	−0.0480	0.6963/0.0013 (significant)

The southern region serves as the primary pathway for exogenous heat input, with its relatively high SST sustained by heat transport associated with the residual YSWC. The western region, encompassing the DC section, is shallow and adjacent to the continent, where winter monsoon-induced cooling and vertical mixing are most intense, facilitating the formation and sinking of low-temperature water. Situated between these two regions, the central region functions as a transitional zone; its relatively deep waters enable the maintenance of a certain horizontal temperature gradient under the strong mixing, making it a convergence zone of southern warm water and western cold water. Consequently, the SST in this central region reflects a dynamic thermodynamic equilibrium between the opposing thermal influences.

Furthermore, the central region exhibits the highest correlation with the cold center (CC = 0.8356), compared with the western region where the DC transect is located (CC = 0.7646) and the southern region to the south of the Chengshantou–Changshanchuan transect (CC = 0.8121). The maximum and minimum February SST in the central region occurred in 2020 and 2011, respectively, aligning with the corresponding extrema of the cold center, rather than the 2004 and 2018 occurrences in the entire domain. Notably, when SST in the central region is relatively high ($\geq$5 °C), the BMT is also elevated, as seen in 2004, 2007, 2009, 2017, 2019, and 2020. Based on these observations, we preliminarily infer that the February SST in the central region may serve as the primary source region of the cold water mass.

The clustering results above capture the overall spatiotemporal characteristics of February SST across different regions of the NYS. However, discrepancies remain between these patterns and the variations of the cold center. For instance, the mean SST in the central region increased in 2016, while the BMT at the cold center decreased. This indicates that the intensity of the cold-water core is not solely governed by the regional mean SST, but likely also influenced by day-to-day SST variability, atmospheric cooling, and local ocean circulation. Therefore, it is necessary to further investigate the daily evolution of SST during February. In addition, considering that the cold center is also significantly correlated with the southern region and is primarily located near the DC section in the western region, the following analysis focuses on how short-term SST variations in the central, southern, and western regions during February influence the intensity of the NYSCWM.

A multi-year clustering analysis was performed for each calendar day using SST data from 1 to 28 February during 2003–2023 (29 February was excluded due to the limited number of leap years). For each day (e.g., 1 February), SST fields across 18 years were used as clustering inputs to identify representative and stable spatial patterns. The clustering

results are shown in Figure 6. Overall, the daily spatial variations within each cluster are relatively minor (Figure 7), indicating that the multi-year daily SST distributions in February exhibit strong spatiotemporal consistency. This provides a robust quantitative foundation for analyzing the background conditions that influence the formation of the NYSCWM in summer.

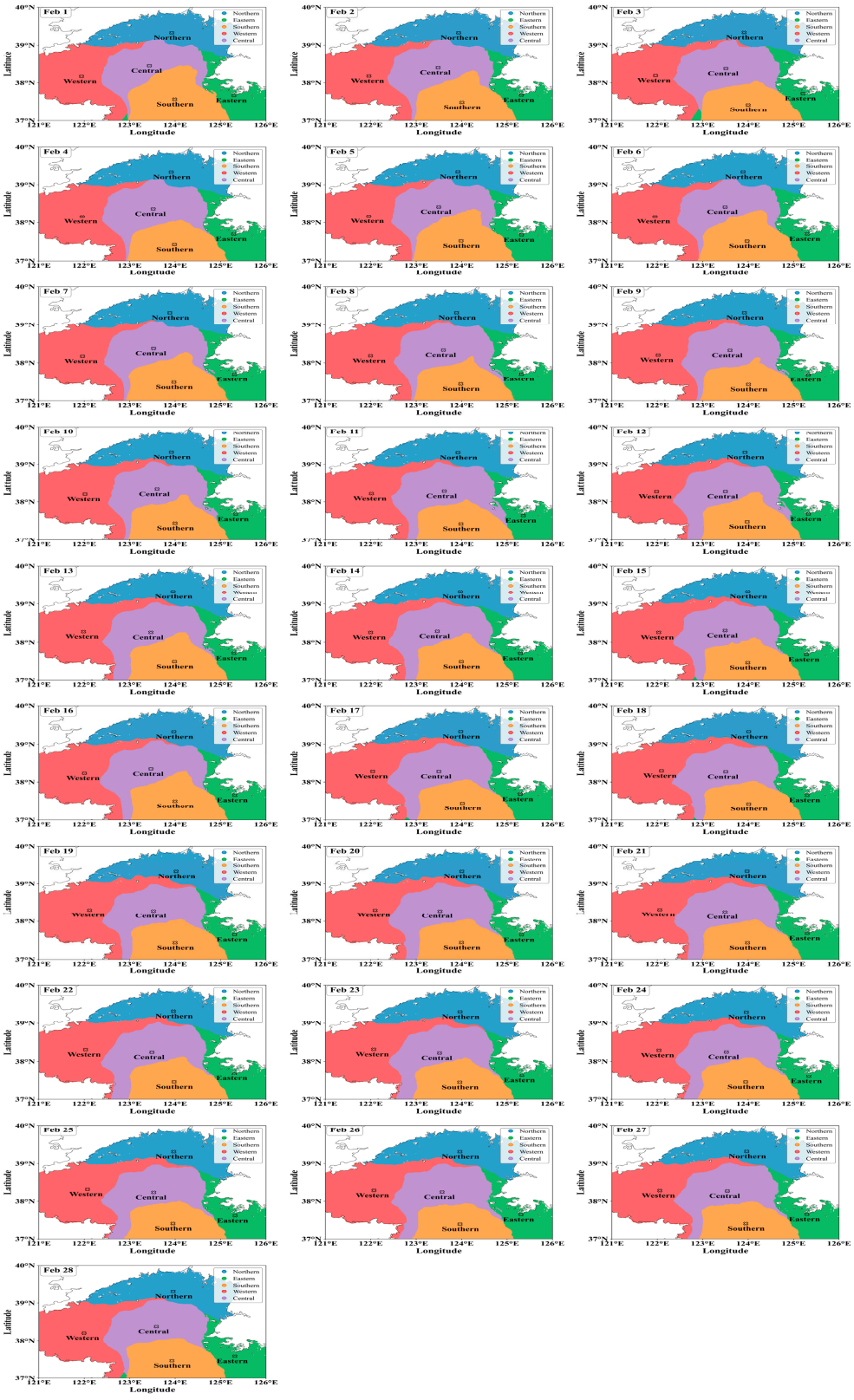

Figure 6. Spatial pattern of multi-year daily clustering results. Cluster centroids are denoted by 'X' markers representing the geometric mean position of each cluster.

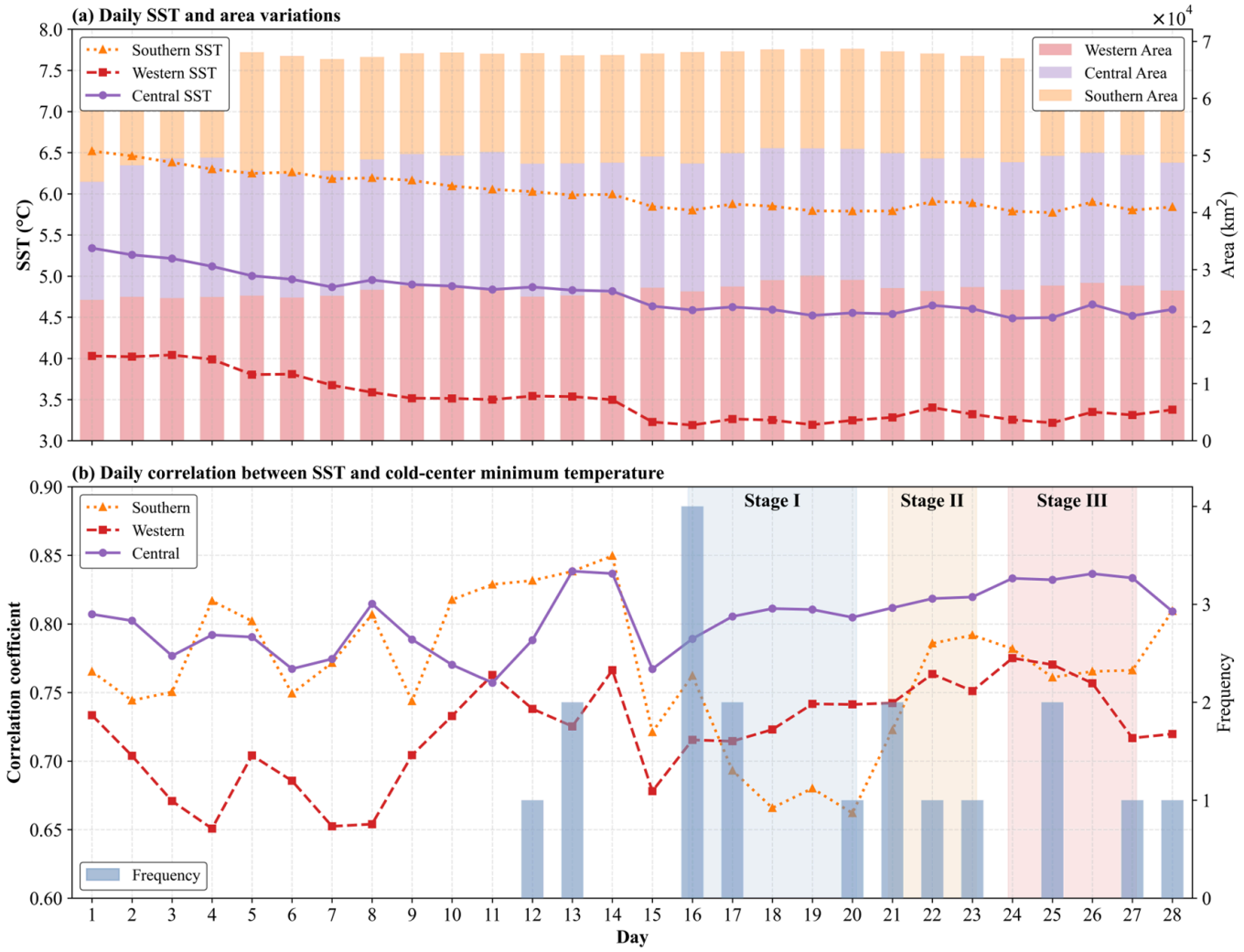

Figure 7. Multi-year daily SST variations in different regions of the NYS and their relationships with the cold center. (**a**) Spatial variations and time-series characteristics of multi-year daily SST in each region; (**b**) correlations between multi-year daily SST in each region and the cold center, and the occurrence frequency of dates corresponding to the minimum SST in the central region in February.

Figure 7 shows the February SST variations in the central, southern, and western regions, together with their correlations with the BMT of the cold center. Across all three regions, SST exhibits a gradual cooling trend through February, with a pronounced decline around 15 February. During the latter half of the month, the SST remains at its lowest levels, indicating a critical period for the formation of local cold water, as shown in Figure 7a. Correspondingly, Figure 7b shows a sharp drop in correlation on 15 February across all three regions, likely associated with this rapid cooling event, followed by a gradual recovery thereafter. Thereafter, the correlations gradually recover. Compared with the western and southern regions, the central region displays the highest correlations and a more stable upward trend. These results suggest that SST variations in the central region during the second half of February play a key spatiotemporal role in modulating the intensity of the cold water mass.

To further identify the timing and characteristics of low-temperature events, the dates corresponding to the minimum February SST in the central region for each year were analyzed, as shown by the bar chart in Figure 7b. The results show that, over the 18-year period, the minimum SST consistently occurred in mid-to-late February, with most events concentrated between 16 and 28 February. Notably, among these, 16 February exhibited

the highest frequency, appearing four times, followed by 13, 17, 21, and 25 February, each occurring twice.

In summary, during winter, the combined effects of strong surface cooling and wind-induced mixing promote vigorous vertical homogenization of the water column, resulting in the lowest SSTs in the central NYS by late February. This period is characterized by both a high correlation with BMT (CC = 0.8396, $p < 0.0001$) and a stable spatial pattern. Combined with the established understanding that the YSCWM originates from the residual water remaining after winter cooling and vertical mixing [1], these findings further suggest that the SST in the central region during the second half of February serves as the primary source for the formation of the YSCWM.

4. Model Fitting and Validation

Based on the results of this study, the SST in the second half of February serves as an important precursor signal for the formation of the cold water mass. To further verify the spatial source regions that influence variations in cold water mass intensity, the interannual SST variations in the central, western, and southern regions were analyzed (as indicated by the solid lines in Figure 8), and two types of regression models were constructed: three univariate models based on SST in each region and a multivariate model integrating SST from all three regions. As shown by the dashed lines in Figure 8, all four fitted curves broadly reproduce the temporal evolution of the cold center, and the multivariate model exhibits the highest correlation, although a relatively large error remains (RMSE = 0.4084). None of the models successfully captures the observed in 2020. Among the univariate models, the central-region model performs better than those of the western and southern regions. Moreover, the SST in the central region during the second half of February shows the strongest correlation with BMT (CC = 0.8396), indicating a high degree of coherence between SST variability in this region and the interannual fluctuations in NYSCWM intensity.

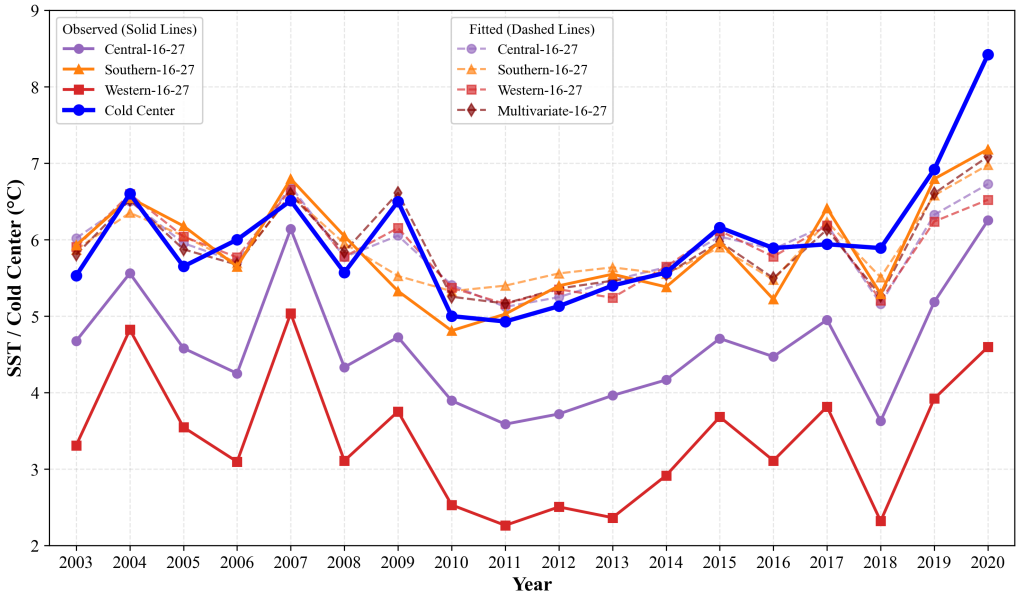

Figure 8. Comparison of observed and fitted SST time series between regional averages and the cold center. Solid lines represent observations, while dashed lines represent fitted values.

Building on these findings, this study further examines the SST evolution in the central region. Considering that SST remains persistently low during the second half of February, which represents a critical period most sensitive to cold water mass formation, to avoid

the limitation of single-date SST data in accurately reflecting the mixing, stabilization, and sinking processes of cold water, this study selected a continuous period from 16 to 27 February as the key analysis window. Based on the temporal variations in SST and the corresponding correlation analysis (Figure 7), this period is divided into three stages. The first stage (16–20 February) encompasses the initial response period, during which the frequency of the minimum SST in the central region was highest, and the correlation with the NYSCWM gradually increased. The second stage (21–23 February) encompasses the response enhancement period, characterized by persistently high frequencies of minimum SST occurrence and a continuous strengthening of correlation. The third stage (24–27 February) encompasses the peak stable period, when the SST remains in a stable low-temperature state and the correlation with the NYSCWM reaches its maximum.

To further evaluate the influence of SST in the central region on the NYSCWM's intensity, regression equations were developed using SST from the three identified phases to predict the BMT of the cold center. Specifically, the SST data of the central region in February and the BMT in August from 2003 to 2019 were used as training samples, while the data from 2020 served as independent test samples. Given the nonlinear characteristics of SST, a quadratic polynomial regression model was constructed, and the regression coefficients were estimated using the least-squares method, as expressed in Equation (3).

$$y = 2.7486x_1^2 + 1.8955x_2^2 - 1.0445x_3^2 - 5.5921x_1x_2 - 2.4117x_1x_3 + 4.7026x_2x_3 + 11.7485x_1 - 12.4060x_2 - 1.2862x_3 + 8.452 \quad (3)$$

In the equation, y is the dependent variable (predicted BMT), and x_1, x_2, x_3 represent the independent variables (the observed SSTs in the central region during February). Specifically, x_1 corresponds to the mean SST from 16 to 20 February, x_2 to the mean SST from 21 to 23 February, and x_3 to the mean SST from 24 to 27 February.

Based on the training samples from 2003 to 2019, the fitting results are shown in Figure 9. In Figure 9a, the blue solid and green dashed lines represent the observed and predicted BMT time series, respectively, while Figure 9b shows the corresponding scatter plot. The predicted values showed a strong correlation with the observations (CC = 0.9146), with a low mean absolute error (MAE) of approximately 0.1820 and an explained variance of 83.64% (R^2 = 0.8364). The significance test of the model (F-test) yields F = 3.977 and p = 0.0412. These results indicate that the regression equation effectively captures the interannual variability of BMT and closely reproduces the overall trend of the observations. Notably, the model successfully reproduces both the lowest BMT in 2011 and the highest BMT in 2019. Furthermore, as shown in Figure 9b, most data points lie close to the 1:1 reference line and fall within the 95% confidence interval, further confirming the model's reliability in reproducing variations in the NYSCWM.

Based on the test dataset, the SST values of the central region during the second half of February 2020 ($x_1 = 6.0889, x_2 = 6.2005, x_3 = 6.4939$) were substituted into Equation (3) to obtain the predicted BMT in August 2020, which was approximately 8.31 °C. This prediction deviates only slightly from the observed value (8.42 °C), indicating high predictive accuracy. These findings further support the conclusion that February SST in the central region serves as the primary source signal for the formation of the NYSCWM and provides a scientific basis for understanding its interannual variability.

Similarly, using SST values from the central region during the second half of February for 2021–2025, the corresponding August BMT can also be predicted, as shown by the orange curve in Figure 9a. Because observed BMT data after 2021 are not yet available, the February SST of the central region was used as a reference metric. As illustrated in Figure 9, the predicted BMT for 2020–2025 shows strong coherence with the interannual variations of central-region SST, with a correlation coefficient of approximately 0.9616 (p = 0.0022), and

both exhibit a modest decreasing trend. This indicates that the model is capable not only of capturing the intense phases of the cold water mass but also of faithfully representing its weakening characteristics.

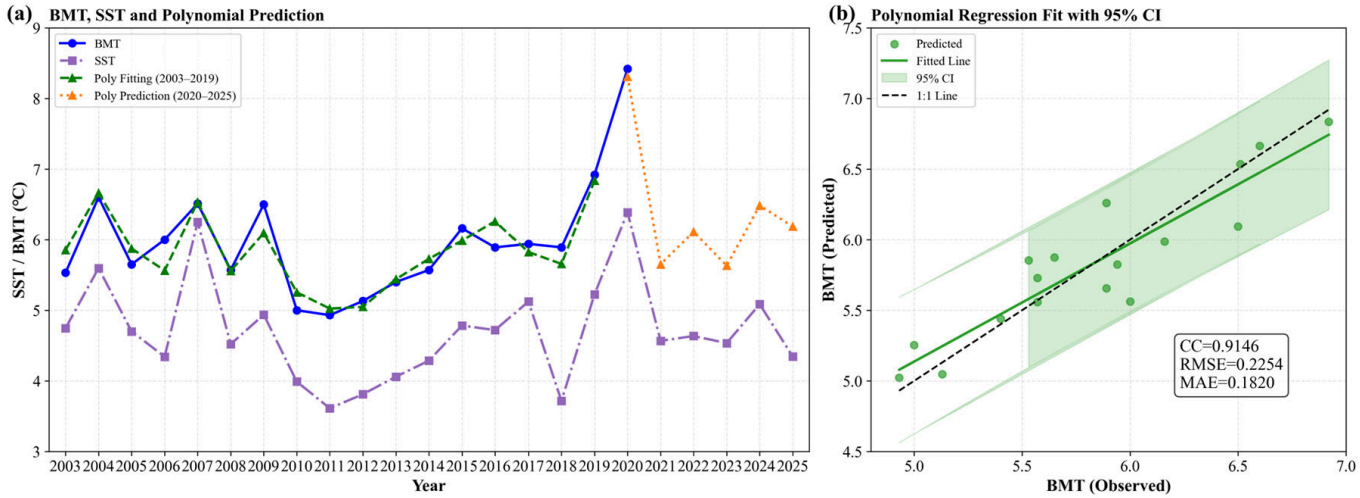

Figure 9. Comparison of fitted and predicted August BMT during 2003–2019 and 2020–2025. (**a**) The blue curve denotes the observed BMT at the cold center, the purple curve denotes the observed SST, the green curve represents the quadratic polynomial regression fit to the August BMT (2003–2019), and the orange curve indicates the predicted August BMT (2020–2025). (**b**) Scatter plot showing the fitted relationship between the observed Cold Center and the quadratic polynomial regression model; the black dashed line denotes the 1:1 reference line, and the shaded area indicates the 95% confidence interval.

Overall, the quadratic regression model effectively captures the quantitative relationship between BMT and SST in the central region during the second half of February, providing a reliable approach for predicting the intensity of the NYSCWM. The results of this study enable prediction of NYSCWM intensity up to six months in advance and provide a robust technical basis for understanding its interannual variability.

5. Conclusions and Future Work

Based on historical observations, cruise data, and high-resolution sea surface temperature datasets, this study investigated the interannual variations in the summer bottom-layer minimum temperature of the Northern Yellow Sea Cold Water Mass, quantified its relationship with wintertime sea surface temperature in the Northern Yellow Sea, and focused on revealing the local core regions and key temporal windows that govern the variability of cold-center intensity. The main findings are summarized as follows:

(1) From 2003 to 2020, the bottom-layer minimum temperature of the cold center in August shows a significant overall warming trend, with an interannual rate of approximately 0.0533 °C yr^{-1}. Specifically, it exhibited a high-frequency oscillatory cooling trend during 2003–2011, followed by a sharp warming trend of 0.2862 °C yr^{-1} during 2011–2020. This indicates that the intensity of the Northern Yellow Sea Cold Water Mass is gradually weakening. This warming trend is generally consistent with previous studies reporting long-term warming and weakening trends of the Yellow Sea Cold Water Mass.

(2) The sea surface temperature in the central region during the second half of February shows a significant positive correlation with the August bottom-layer minimum temperature (Correlation Coefficient = 0.8396). This value is notably higher than the correlation for the February mean sea surface temperature of the entire study area (Correlation Coefficient = 0.6963), the western sea area along the Dalian–Chengshantou transect with a

correlation coefficient of 0.7646, and the southern sea area south of the Chengshantou–Changshanchuan transect with a correlation coefficient of 0.8121. Therefore, we infer that sea surface temperature in the central Northern Yellow Sea during the second half of February may be the dominant factor affecting the low-temperature water of the Northern Yellow Sea Cold Water Mass.

(3) The K-means clustering analysis shows that the key spatiotemporal characteristics of the dominant factors are as follows. (a) Spatially, the central region of the Northern Yellow Sea is a convergence zone between the southern warm water (originating from the central Yellow Sea thermal input) and the western cold water (adjacent to the northern Shandong Peninsula and the southern coast of the Liaodong Peninsula). This region plays a crucial role in the dynamic thermal equilibrium and exhibits the strongest correlation with the bottom-layer minimum temperature, along with the most stable variation pattern. (b) Temporally, the second half of February represents the critical period influencing the cold water mass intensity. This period can be divided into three phases: initial response (16–20 February), enhancement response (21–23 February), and peak stability (24–27 February). During this period, the sea surface temperature in the central region remains at a persistently low level, indicating a stable thermal state that supports the formation and development of the cold water mass.

(4) A multiple regression model was constructed, effectively quantifying the relationship between the February sea surface temperature in the central Northern Yellow Sea across its three temporal phases and cold water mass intensity. This further verifies the rationality of the above inference. That is, sea surface temperature in the central region during the second half of February is a key signal for the formation and intensity variation of the cold water mass. The results of this study can not only predict the cold water mass's bottom-layer minimum temperature in advance, but also provide a scientific basis for understanding its interannual evolution and for informing marine environmental management and economic development.

This study mainly focuses on the relationship between winter sea surface temperature and summer bottom-layer minimum temperature in the Northern Yellow Sea and does not explicitly include other relevant processes, such as atmospheric forcing, air–sea heat flux, and ocean circulation. These factors determine the distribution and stability of winter sea surface temperature, consequently affecting the formation and evolution of the summer cold water mass. Although the present work is conducted for the Northern Yellow Sea Cold Water Mass, the analytical framework and research method adopted here are generalizable and can be extended to other marginal seas and continental shelf regions worldwide, such as the Seto Inland Sea, Japan. Moreover, previous studies have indicated that the Northern Yellow Sea Cold Water Mass is also closely linked to springtime warming and stratification processes. As the present work only characterized the quantitative relationship between February sea surface temperature and cold water intensity, it did not address the regulatory effects of spring sea surface temperature, thermocline thickness, or stratification strength. Therefore, future studies will comprehensively consider factors such as springtime sea surface temperature and air–sea heat transfer processes. Such efforts will provide a more robust physical basis for predicting the interannual variability of the Northern Yellow Sea Cold Water Mass and further enhance forecasting accuracy and the management efficiency of the regional marine environment.

Author Contributions: Conceptualization, X.C.; methodology, X.C.; software, X.C. and Z.M.; validation, Z.M. and M.S.; formal analysis, X.C.; investigation, Z.M., M.S. and J.S.; resources, Z.M. and T.L.; data curation, Z.M. and T.L.; writing—original draft preparation, X.C. and Z.M.; writing—review and editing, X.C., Z.L. and Y.L.; visualization, X.C., Z.M. and T.L.; supervision, X.C., Z.L. and Y.L.; project

administration, Z.M.; funding acquisition, X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 42306218) and the Natural Science Foundation of Hebei Province, China (No. F2023407003).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. He, C.B.; Wang, Y.X.; Lei, Z.Y.; Xu, S. Preliminary study on the formation and properties of the Yellow Sea Cold Water Mass. *Oceanol. Limnol. Sin.* **1959**, *1*, 11–15.
2. Wei, C.J.; Tang, X.H.; Ge, K.; Xu, A.Q.; Li, Y.L.; Jiang, Y.; Rong, Z.R.; Yu, F. Observed seasonal evolution and origins of the western Yellow Sea Cold Water Mass. *Front. Mar. Sci.* **2025**, *12*, 1556069. [CrossRef]
3. Li, X.; Wang, X.; Chu, P.C.; Zhao, D.L. Low-Frequency Variability of the Yellow Sea Cold Water Mass Identified from the China Coastal Waters and Adjacent Seas Reanalysis. *Adv. Meteorol.* **2015**, *2015*, 269859. [CrossRef]
4. Li, J.C.; Li, G.X.; Xu, J.S.; Dong, P.; Qiao, L.L.; Liu, S.D.; Sun, P.K.; Fan, Z.S. Seasonal evolution of the Yellow Sea Cold Water Mass and its interactions with ambient hydrodynamic system. *J. Geophys. Res. Ocean.* **2016**, *121*, 6779–6792.
5. Liu, X.; Yu, F.; Chen, Z.F.; Si, G.C.; Nan, F.; Wang, J.F.; Ren, Q.; Hu, Y.B. The Critical Role of Thermal Stratification Associated with the Yellow Sea Cold Water Mass in Modulating Winter Sea Surface Temperature. *J. Geophys. Res. Ocean.* **2024**, *129*, e2023JC020373. [CrossRef]
6. Su, J.L.; Huang, D.J. On the Current Field Associated with the Yellow Sea Cold Water Mass. *Oceanol. Limnol. Sin.* **1995**, *S1*, 1–7.
7. Dong, S.L. Researching Progresses and Prospects in Large Salmonidae Farming in Cold Water Mass of Yellow Sea. *Period. Ocean Univ. China* **2019**, *49*, 1–6.
8. Wu, R.; Li, J.C.; Ye, Z.J.; Wang, B.; Liu, S.D.; Dong, X.Q.; Tian, Y.J. Growth and Distribution of Young Pacific Cod in Yellow Sea. *Period. Ocean Univ. China* **2020**, *50*, 63–73.
9. Weng, X.C.; Zhang, Y.K.; Wang, C.M.; Zhang, Q.L. The Variational Characteristics of the Huanghai Sea (Yellow Sea) Cold Water Mass. *Oceanol. Limnol. Sin.* **1989**, *19*, 119–131.
10. Yu, F.; Zhang, Z.X.; Diao, X.Y.; Guo, J.S.; Tang, Y.X. Analysis of evolution of the Huanghai Sea Cold Water Mass and its relationship with adjacent water masses. *Haiyang Xuebao* **2006**, *28*, 26–34.
11. Zhang, S.W.; Wang, Q.Y.; Lü, Y.; Cui, H.; Yuan, Y.L. Observation of the seasonal evolution of the Yellow Sea Cold Water Mass in 1996–1998. *Cont. Shelf Res.* **2007**, *28*, 442–457. [CrossRef]
12. Zhang, Y.K.; Yang, Y.L. Analyses of the Variational Characteristics of the North Huanghai Sea Cold Water Mass. *Mar. Forecast.* **1996**, *4*, 16–22.
13. Liu, C.H.; Chen, X.; Wang, C.X.; Liu, Z.L.; Jia, S.Y.; Wang, X. Review on seasonal and interannual variation characteristics of low temperature center in the Yellow Sea cold water mass. *Mar. Environ. Sci.* **2024**, *43*, 475–488.
14. Yang, Y.; Li, K.P.; Du, J.T.; Liu, Y.L.; Liu, L.; Wang, H.W.; Yu, W.D. Revealing the Subsurface Yellow Sea Cold Water Mass from Satellite Data Associated with Typhoon Muifa. *J. Geophys. Res. Ocean.* **2019**, *124*, 7135–7152. [CrossRef]
15. Liu, X.C.; Zhai, F.G.; Yan, J.J.; Gu, Y.Z.; Wang, Y.C.; Li, P.L.; Wu, K.J. Three-Dimensional Temperature Responses to Northward-Moving Typhoons in the Shallow Stratified Yellow Sea in Summer. *J. Geophys. Res. Ocean.* **2022**, *127*, e2022JC019091. [CrossRef]
16. Yin, J.H.; Zhang, G.T.; Li, C.L.; Wang, S.W.; Wan, A.Y. Annual variation in abundance of Saggita crassa and its relationship with environment conditions around the Zhangzi Island, Northern Yellow Sea. *Haiyang Xuebao* **2016**, *38*, 86–94.
17. Xia, Y.Y.; Zhang, J.H.; Liu, Y. Behavioral characteristics and physiological responses to hypoxic stress in *Patinopecten yessoensis*. *J. Fish. Sci. China* **2021**, *28*, 1319–1328.
18. Sun, X.Y.; Gao, X.L.; Zhao, J.M.; Xing, Q.G.; Liu, Y.L.; Xie, L.; Wang, Y.J.; Wang, B.; Lv, J.S. Promoting effect of raft-raised scallop culture on the formation of coastal hypoxia. *Environ. Res.* **2023**, *228*, 115810. [CrossRef] [PubMed]
19. Song, X.; Lin, X.P.; Wang, Y. The Preliminary Study of Long-term Variability of the Yellow Sea Cold Water in Summer and its Possible Reasons. *J. Guangdong Ocean Univ.* **2009**, *29*, 59–63.
20. Park, S.; Chu, P.C.; Lee, J.-H. Interannual-to-interdecadal variability of the Yellow Sea Cold Water Mass in 1967–2008: Characteristics and seasonal forcings. *J. Mar. Syst.* **2011**, *87*, 177–193. [CrossRef]
21. Li, A.; Yu, F.; Diao, X.Y.; Si, G.C. Interannual variability of temperature of the northern Yellow Sea Cold Water Mass. *Haiyang Xuebao* **2015**, *37*, 30–42.
22. Guo, Y.X.; Mo, D.X.; Hou, Y.J. Interannual to Interdecadal Variability of the Southern Yellow Sea Cold Water Mass and Establishment of "Forcing Mechanism Bridge". *J. Mar. Sci. Eng.* **2021**, *9*, 1316. [CrossRef]

23. Yang, J.; Liu, C.L.; Sun, Q.W.; Zhai, L.; Sun, Q.M.; Li, S.J.; Ai, L.B.; Li, X. Interannual Variability and Long-Term Trends in Intensity of the Yellow Sea Cold Water Mass during 1993–2019. *J. Mar. Sci. Eng.* **2023**, *11*, 1888. [CrossRef]
24. Shen, X.Y.; Yao, Z.G.; Bao, X.W.; Li, X.B.; Ding, Y. Interannual Temperature Variations of the Yellow Sea Cold Water Masses: A Comprehensive Analysis from 1976 to 1999. *J. Ocean Univ. China* **2025**, *24*, 875–885. [CrossRef]
25. Jiang, B.J.; Bao, X.W.; Wu, D.X.; Xu, J.P. Interannual variation of temperature and salinity of northern Huanghai Sea Cold Water Mass and its probable cause. *Haiyang Xuebao* **2007**, *29*, 1–10.
26. Guan, B.X. Preliminary Analysis of Water Temperature Variation and Circulation Characteristics of the Yellow Sea Cold Water Mass. *Oceanol. Limnol. Sin.* **1963**, *5*, 255–284.
27. Sun, X.P. Some Examples in Variability of the Cold Water Mass of the Yellow Sea and Sea Ice of the Bohai Caused by the Climatical Anomaly. *Trans. Oceanol. Limnol.* **1980**, *1*, 1–8.
28. Du, B.; Zhang, Y.J.; Shan, Y.C.; Wang, H. The Characteristics of Cold Water Mass Variation at the Bottom of the North Yellow Sea and Its Hydrological Effects on the Mortality of Shellfish Cultured in the Waters of Outer Chang-shan Islands. *Mar. Sci. Bull.* **1996**, *4*, 17–28.
29. Yang, H.W.; Cho, Y.K.; Seo, G.H.; You, S.H.; Seo, J.W. Interannual variation of the southern limit in the Yellow Sea Bottom Cold Water and its causes. *J. Mar. Syst.* **2014**, *139*, 119–127. [CrossRef]
30. Chu, Q.Q.; Yu, H.M.; Li, S.L.; Yu, H.Q.; Ge, J.J. Study of the correlation between SST and inter-annual variation of the cold water mass in the Yellow Sea. *Mar. Forecast.* **2021**, *38*, 21–30.
31. Li, H.; Zhai, F.G.; Dong, Y.J.; Liu, Z.Z.; Gu, Y.Z.; Bai, P. Interannual-decadal variations in the Yellow Sea Cold Water Mass in summer during 1958–2016 using an eddy-resolving hindcast simulation based on OFES2. *Cont. Shelf Res.* **2024**, *275*, 105223. [CrossRef]
32. Zhu, J.Y.; Jie, S.; Guo, X.Y.; Gao, H.W.; Yao, X.H. Air-sea heat flux control on the Yellow Sea Cold Water Mass intensity and implications for its prediction. *Cont. Shelf Res.* **2018**, *152*, 14–26. [CrossRef]
33. Chin, T.M.; Vazquez-Cuervo, J.; Armstrong, E.M. A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sens. Environ.* **2017**, *200*, 154–169. [CrossRef]
34. Chen, X.; Liu, T.; Song, M.; Ma, Z.; Liu, Z. A Long-term forecasting model for sea surface temperature based on wavelet decomposition and adaptive fusion. *J. Sea Res.* **2026**, *211*, 102696. [CrossRef]
35. Agarwal, P.K.; Mustafa, N.H. k-means projective clustering. In Proceedings of the ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS), Paris, France, 13–15 June 2004.
36. Ruela, R.; Sousa, M.C.; de Castro, M.; Dias, J.M. Global and regional evolution of sea surface temperature under climate change. *Glob. Planet. Change* **2020**, *190*, 103190. [CrossRef]
37. Jin, S.S.; Nie, X.W.; Wang, G.L.; Teng, F.; Xu, T.F. Analysis of the Distribution and Seasonal Variability of the South China Sea Water Masses Based on the K-means Cluster Method. *J. Mar. Sci. Eng.* **2023**, *11*, 485. [CrossRef]
38. Draper, N.R.; Smith, H. *Applied Regression Analysis*, 3rd ed.; Wiley: New York, NY, USA, 1998; pp. 1–713.
39. Furner, R.; Haynes, P.; Munday, D.; Paige, B.; Jones, D.C.; Shuckburgh, E. A sensitivity analysis of a regression model of ocean temperature. *Environ. Data Sci.* **2022**, *1*, e11. [CrossRef]
40. Chang, M.H.; Huang, Y.C.; Cheng, Y.H.; Terng, C.T.; Chen, J.; Jan, J.C. Revisiting regression methods for estimating long-term trends in sea surface temperature. *Nat. Hazards Earth Syst. Sci.* **2024**, *24*, 2481–2494. [CrossRef]