Regional Studies in Marine Science 103 (2026) 105514

Contents lists available at ScienceDirect

Regional Studies in Marine Science

journal homepage: www.elsevier.com/locate/rsma

Endpoint-separated modelling of environmental associations with *Noctiluca scintillans* occurrence probability, conditional positive abundance, and Bloom90 probability in the southern Black Sea

İlknur Yıldız [a,*,1], Ömerhan Dürrani [b,2]

[a] Institute of Marine Sciences and Technology, Karadeniz Technical University, Trabzon, 61080, Türkiye
[b] Department of Molecular Biology and Genetics, Faculty of Science, Karadeniz Technical University, Trabzon, 61080, Türkiye

ARTICLE INFO

Keywords:
Blocked validation
Cluster-aware inference
Environmental gradients
Heterotrophic dinoflagellate
Monitoring prioritisation
Upper-tail abundance
Türkiye

ABSTRACT

Noctiluca scintillans (Macartney) Kofoid & Swezy is a heterotrophic bloom-forming dinoflagellate whose dense aggregations may alter trophic interactions and nutrient cycling. This study examined environmental and temporal associations of *N. scintillans* using a seven-year fixed-station monitoring dataset covering winter and summer sampling at 21 stations from İğneada to Hopa along the southern Black Sea coast during 2018–2024. An endpoint-separated modelling framework distinguished occurrence probability, conditional positive abundance, and Bloom90 probability. Bloom90 was defined as a dataset-specific upper-decile abundance endpoint corresponding to ~1744 ind. m^{-3}. *Noctiluca scintillans* was detected in 211 of 292 valid abundance observations (72.3%) and reached a maximum abundance of 66,751 ind. m^{-3}. Each 6.88°C increase in temperature, equivalent to one complete-dataset standard deviation, was associated with a 91.0% reduction in occurrence odds (odds ratio = 0.090, 95% confidence interval 0.031–0.223). The occurrence-probability model retained good discrimination under leave-one-station-out blocked validation (area under the receiver operating characteristic curve [ROC AUC] = 0.860) but lower discrimination under leave-one-year-out blocked validation (ROC AUC = 0.780). Conditional positive abundance varied principally among years and showed limited out-of-sample performance. After environmental adjustment, summer sampling events had 13.7-fold higher Bloom90 odds than winter sampling events. In the complete dataset, Bloom90 probability was negatively associated with temperature and phosphate phosphorus and positively associated with Chlorophyll-a; the Chlorophyll-a association was not retained after exclusion of 2021. The Bloom90-probability model showed modest discrimination under leave-one-station-out blocked validation (ROC AUC = 0.698) and no meaningful discrimination under leave-one-year-out blocked validation (ROC AUC = 0.517). S2/K0 in the Istanbul Strait outflow sector had the highest observed Bloom90 frequency, but wide, overlapping station-level confidence intervals precluded precise spatial ranking. The endpoint-separated modelling framework therefore provides an explanatory basis for monitoring prioritisation but does not support temporally transferable forecasting of Bloom90 events.

1. Introduction

The Black Sea is a semi-enclosed, eutrophication-sensitive basin characterised by pronounced hydrographic and biogeochemical gradients, including oxygen limitation (Oguz and Velikova, 2010; Alkan et al., 2022; Yücel et al., 2026). Its planktonic food-web structure is further influenced by nutrient enrichment, gelatinous-plankton interactions, and physical–biological coupling (Shiganova et al., 2008; Mikaelyan et al., 2014). The red heterotrophic form of *Noctiluca scintillans* (Macartney) Kofoid & Swezy is a recurrent bloom-forming dinoflagellate whose population dynamics reflect phytoplankton production, nutrient recycling, hydrographic stability, and physical aggregation (Elbrächter and Qi, 1998; Miyaguchi et al., 2006; Harrison et al., 2011; Kordubel et al., 2024). Although not generally considered directly toxic, dense aggregations may restructure plankton communities, enhance nutrient regeneration, reduce dissolved oxygen, and redirect carbon and

* Corresponding author.
E-mail addresses: ilknurk@ktu.edu.tr (İ. Yıldız), omerhandurrani@ktu.edu.tr (Ö. Dürrani).
[1] 0000-0003-2424-8644
[2] 0000-0003-1775-8662

https://doi.org/10.1016/j.rsma.2026.105514

Available online 7 October 2026
2352-4855/© 2026 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

nutrients towards heterotrophic recycling and microbial processing (Elbrächter and Qi, 1998; Harrison et al., 2011; Ara et al., 2013; Drits et al., 2013; Baliarsingh et al., 2016; Kordubel et al., 2024).

Within the Black Sea, *N. scintillans* abundance varies with seasonal production and trophic structure (Nikishina et al., 2011; Mikaelyan et al., 2014), against a regional background of long-term temperature, salinity, and Chlorophyll-a variability (Alkan et al., 2013), eutrophication (Porumb, 1992), and gelatinous-plankton dynamics (Kovalev and Piontkovski, 1998). In the south-eastern Black Sea, maximum abundance was recorded in April, and abundance was reported to be negatively associated with temperature (Şen Ozdemir et al., 2017). An independent historical event off Rize occurred at 9.9–11.7°C under stable surface conditions, low phosphate phosphorus (PO_4-P), and elevated Chlorophyll-a; abundant diatoms and *Melosira* spp. were also observed within *N. scintillans* cells (Kopuz et al., 2014). This event documented the co-occurrence of cool conditions, a bulk phytoplankton/productivity-related signal, phosphate depletion, and hydrographic stability but did not resolve the processes governing longer-term abundance patterns.

The heterotrophic ecology of *N. scintillans* differentiates its environmental associations from those of autotrophic bloom-forming microalgae (Elbrächter and Qi, 1998; Harrison et al., 2011). As a heterotroph, *N. scintillans* does not derive production directly from dissolved inorganic nutrients; observed nutrient associations may instead reflect antecedent phytoplankton development or particulate food availability (Elbrächter and Qi, 1998; Kiørboe and Titelman, 1998; Nakamura, 1998; Zhang et al., 2016; Sheng et al., 2022). Nutrient regeneration through feeding, excretion, and bloom decay may further modify contemporaneous concentrations (Drits et al., 2013; Baliarsingh et al., 2016). Elevated Chlorophyll-a may represent greater bulk phytoplankton biomass or productivity, whereas low PO_4-P may be consistent with phosphate-depleted post-production conditions rather than direct nutrient inhibition (Baliarsingh et al., 2016; Ollevier et al., 2021; Sheng et al., 2022). Chlorophyll-a nevertheless remains a bulk phytoplankton/productivity-related proxy and does not directly quantify edible prey biomass, prey quality, or taxonomic composition.

Local abundance may also be modified by transport and physical aggregation (Huang and Qi, 1997; Dela-Cruz et al., 2003; Miyaguchi et al., 2006). Positive buoyancy and surface aggregation make *N. scintillans* responsive to weak turbulence, stratification, wind-driven convergence, coastal accumulation, and retention-prone circulation (Huang and Qi, 1997; Dela-Cruz et al., 2003; Miyaguchi et al., 2006; Kopuz et al., 2014; Kordubel et al., 2024). Along the Turkish coast of the southern Black Sea, the Istanbul Strait outflow, coastal–offshore gradients, and local circulation may generate spatially heterogeneous aggregation conditions, particularly at western stations. Short-lived surface patches may develop when population growth, food availability, and physical concentration coincide (Huang and Qi, 1997; Dela-Cruz et al., 2003; Miyaguchi et al., 2006; Kopuz et al., 2014). Recurrent high-abundance locations may therefore be consistent with candidate retention settings, although hydrodynamic retention cannot be inferred without direct circulation and transport measurements.

Occurrence probability, conditional positive abundance, and upper-tail abundance are rarely evaluated as distinct endpoints. *Noctiluca scintillans* abundance distributions are typically zero-containing, strongly right-skewed, spatially patchy, and characterised by episodic extreme values (Dela-Cruz et al., 2003; Kopuz et al., 2014). Environmental conditions associated with occurrence probability may differ from those associated with conditional positive abundance or Bloom90 probability. A single abundance response may obscure these differences by combining non-detections, low-to-moderate positive observations, and upper-tail events.

This distinction is particularly relevant to the southern Black Sea, where widespread occurrence under cooler conditions may coexist with infrequent upper-tail events during shorter environmentally favourable periods (Kopuz et al., 2014; Mikaelyan et al., 2014; Şen Ozdemir et al., 2017). An endpoint-separated modelling framework permits evaluation of whether temperature, Chlorophyll-a, PO_4-P, season, year, and station-level variation show shared or endpoint-specific associations. Dataset-specific upper-tail events may identify candidate sampling periods for direct measurement of dissolved-oxygen depletion and ammonium enrichment (Baliarsingh et al., 2016; Sathishkumar et al., 2026), ingestion of toxigenic prey and associated phycotoxin-vector potential (Escalera et al., 2007), microbial restructuring (Wang et al., 2025), and bloom-decay chemistry (Sathishkumar et al., 2026). Such event-responsive sampling may be required because conventional monitoring may not resolve short-lived, spatially heterogeneous bloom processes (Kordubel et al., 2024). Model-estimated Bloom90 probability may support monitoring prioritisation but cannot demonstrate prey-mediated phycotoxin transfer, microbial restructuring, oxygen stress, or broader ecological effects.

A seven-year fixed-station dataset from the Turkish coast of the southern Black Sea was therefore used to distinguish three non-equivalent endpoints: occurrence probability, conditional positive abundance, and Bloom90 probability. Bloom90 was defined as a dataset-specific upper-decile abundance endpoint. The objectives were to: (i) quantify seasonal, interannual, and station-level variation in *N. scintillans* occurrence and abundance; (ii) estimate adjusted associations of temperature, Chlorophyll-a, and PO_4-P with occurrence probability, conditional positive abundance, and Bloom90 probability; (iii) identify recurrent high-abundance locations and quantify uncertainty in station-level Bloom90 frequencies; and (iv) assess endpoint-specific robustness, spatial and temporal transportability, and temperature functional form.

2. Materials and methods

2.1. Study area and monitoring design

Data were generated through monitoring conducted under the national Integrated Marine Pollution Monitoring Programme (Denizlerde Bütünleşik Kirlilik İzleme İşi; DEN-İZ), implemented by the TÜBİTAK Marmara Research Center (TÜBİTAK MAM) on behalf of the Ministry of Environment, Urbanization and Climate Change of the Republic of Türkiye. The analytical dataset comprised 294 station–season–year observations collected at 21 fixed monitoring stations during winter and summer from 2018 to 2024 (Table 1; Fig. 1). Each station–season–year combination constituted one observational unit.

The monitoring network extended from İğneada to Hopa and encompassed the Istanbul Strait outflow and the Sakarya, Karadeniz Ereğlisi, Zonguldak, Bartın, İnebolu, Sinop, Samsun, Ordu, Giresun, Trabzon, Arsin, Rize, and Hopa sectors. Station depths ranged from 18.5 to 1530 m and encompassed coastal–offshore and west–east environmental gradients.

Analyses were restricted to upper-layer zooplankton samples collected above the thermocline. Where depth-stratified sampling was conducted, deeper samples were excluded to prevent the combination of non-equivalent pelagic habitats. Thermocline depth was identified from conductivity–temperature–depth profiles, and temperature was averaged across the upper-layer interval corresponding to each zooplankton tow.

2.2. Field sampling, preservation, and environmental measurements

Zooplankton samples were collected by vertical tow using a UNESCO WP2 plankton net with a mouth diameter of 0.57 m and a mesh size of 200 μm. Each tow sampled the upper layer above the thermocline identified from the pre-tow temperature–salinity profile. The net was equipped with a flowmeter, and one tow was conducted per station–season–year observation at the lowest practicable retrieval speed. Filtered seawater volume was calculated as:

Table 1
Fixed monitoring stations included in analyses of *Noctiluca scintillans* along the southern Black Sea coast, 2018–2024.

Station ID	Station code	Province	Monitoring locality	[a]Survey stratum	Latitude (°N)	Longitude (°E)	Depth (m)	Valid abundance observations, *n*
S1	TRK2Y	Kırklareli	İğneada–Danube-influenced sector	DDB1	41.850	28.324	60	14
S2	K0	Istanbul	Istanbul Strait outflow, Line 1	SYB1	41.226	29.134	68	14
S3	1H15	Istanbul	Istanbul Strait outflow channel, Line 1	DDB1	41.454	29.335	1000	13
S4	TRK11	Sakarya	Sakarya River-influenced sector	DDB1	41.167	30.640	47	14
S5	TRKE2	Zonguldak	Karadeniz Ereğlisi sector	SYB3	41.288	31.372	60	14
S6	TRK13	Zonguldak	Zonguldak nearshore sector	SYB3	41.459	31.773	19	14
S7	TRK14Y	Zonguldak	Zonguldak offshore sector	SYB3	41.487	31.759	87	14
S8	TRK21	Bartın	Bartın sector	SYB4	41.698	32.220	130	14
S9	TRK23Y	Kastamonu	İnebolu sector	SYB5	42.016	33.767	70	13
S10	TRK29	Sinop	Sinop-1 outer sector	SYB6	41.909	35.184	47	14
S11	TRK28	Sinop	Sinop-1 nearshore sector	SYB6	42.022	35.156	18.5	14
S12	TRK35	Samsun	Samsun sector	SYB9	41.347	36.391	50	14
S13	TRK44	Ordu	Ordu sector	SYB12	41.021	37.908	51	14
S14	TRK47	Giresun	Giresun offshore sector	SYB12	40.933	38.410	55	14
S15	TRK46	Giresun	Giresun nearshore sector	SYB12	40.923	38.404	30	14
S16	TRK54	Trabzon	Trabzon offshore sector	SYB14	41.031	39.725	103	14
S17	TRK53	Trabzon	Trabzon nearshore sector	SYB14	41.018	39.727	57	14
S18	6H20	Trabzon	Arsin offshore, 20 nmi, Line 6	DDB5	41.277	40.061	1530	14
S19	TRK57	Rize	Rize offshore sector	SYB15	41.057	40.536	92	14
S20	TRK55	Rize	Rize nearshore sector	SYB15	41.035	40.540	26	14
S21	TRK63	Artvin	Hopa sector	SYB16	41.435	41.404	102	14

Note. [a]Survey strata are official identifiers used in the national monitoring design.

$$V_f = A(R_2 - R_1)k, \tag{1}$$

where V_f is the filtered seawater volume (m^3), A is the net-mouth area (m^2), R_1 and R_2 are the pre-deployment and post-retrieval flowmeter readings, respectively, and k is the flowmeter calibration coefficient.

After retrieval, the net was rinsed to transfer retained organisms into the cod end. Samples were transferred to polyvinyl chloride containers and preserved in borax-buffered formaldehyde at a final concentration of 4%.

The environmental matrix comprised temperature (°C), Chlorophyll-a (μg L^{-1}), ammonium nitrogen (NH_4-N; μM), nitrate plus nitrite nitrogen (NO_3+NO_2-N; μM), phosphate phosphorus (PO_4-P; μM), silicate (Si; μM), dissolved oxygen (mg L^{-1}), and Secchi depth (m). Temperature, Chlorophyll-a, nutrient concentrations, and dissolved oxygen were matched to the upper-layer interval corresponding to each zooplankton tow. Secchi depth represented an integrated water-column measurement obtained during the same monitoring event using a 30-cm white Secchi disc in accordance with the national monitoring protocol.

2.3. Laboratory analyses and microscopic enumeration

Chlorophyll-a was determined after filtering 2800 ml of seawater through glass-fibre GF/F filters. Pigments were extracted in 90% acetone and quantified using a Shimadzu UV-1900i ultraviolet–visible spectrophotometer. Nutrient concentrations were measured using a four-channel QuAAtro continuous-flow autoanalyser configured for NH_4-N, NO_3 + NO_2-N, PO_4-P, and silicate. Dissolved-oxygen samples were fixed immediately in the field or aboard the vessel and analysed aboard the vessel using the Winkler method. All procedures followed the standard analytical methods specified in the Integrated Marine Pollution Monitoring guidelines.

Noctiluca scintillans abundance was determined using a Zeiss Discovery V12 stereomicroscope and an Olympus light microscope. Preserved samples were processed by Stempel-pipette subsampling. Counts were extrapolated to the total preserved sample volume and standardised by the flowmeter-derived filtered seawater volume:

$$A_i = \frac{n_i(V_t/V_s)}{V_f}, \tag{2}$$

where A_i is *N. scintillans* abundance in observation i (ind. m^{-3}), n_i is the number of individuals counted, V_s is the analysed subsample volume, V_t is the total preserved sample volume, and V_f is the filtered seawater volume.

2.4. Biological endpoints and analytical samples

The primary biological response was *N. scintillans* abundance (ind. m^{-3}). Two of the 294 monitoring observations lacked valid abundance measurements, leaving 292 valid abundance observations for occurrence-probability and Bloom90-probability analyses. These comprised 81 non-detections and 211 positive observations. Conditional positive-abundance modelling was restricted to the 211 positive observations. Missing abundance values were retained as missing and were not recoded as zero.

Occurrence distinguished detection from non-detection, whereas conditional positive abundance represented abundance magnitude among positive observations. Bloom90 identified observations in the upper decile of the complete-dataset abundance distribution. The unrounded complete-dataset Bloom90 threshold for 2018–2024 was 1744.4 ind. m^{-3} and is reported as ~1744 ind. m^{-3}. Classification used the unrounded threshold; reporting precision therefore did not alter the classification of any observation. Bloom90 represents a dataset-specific upper-decile abundance endpoint rather than a universal ecological, physiological, toxicological, or harmful-bloom threshold.

For observation i, occurrence was defined as:

$$O_i = \begin{cases} 1, & A_i > 0, \\ 0, & A_i = 0, \end{cases} \tag{3}$$

and Bloom90 status as:

$$B_i = \begin{cases} 1, & A_i \geq 1,744.4 \text{ ind. m}^{-3}, \\ 0, & A_i < 1,744.4 \text{ ind. m}^{-3}, \end{cases} \tag{4}$$

where A_i is observed abundance, O_i is the occurrence indicator, and B_i is the Bloom90 indicator.

Stations S3/1H15 and S9/TRK23Y each lacked one valid abundance observation; all other stations contributed 14 valid abundance observations. The primary endpoint-specific models included 292 complete-case observations for abundance, temperature, Chlorophyll-a, and PO_4-P. Ordination analyses incorporating Secchi depth included 280

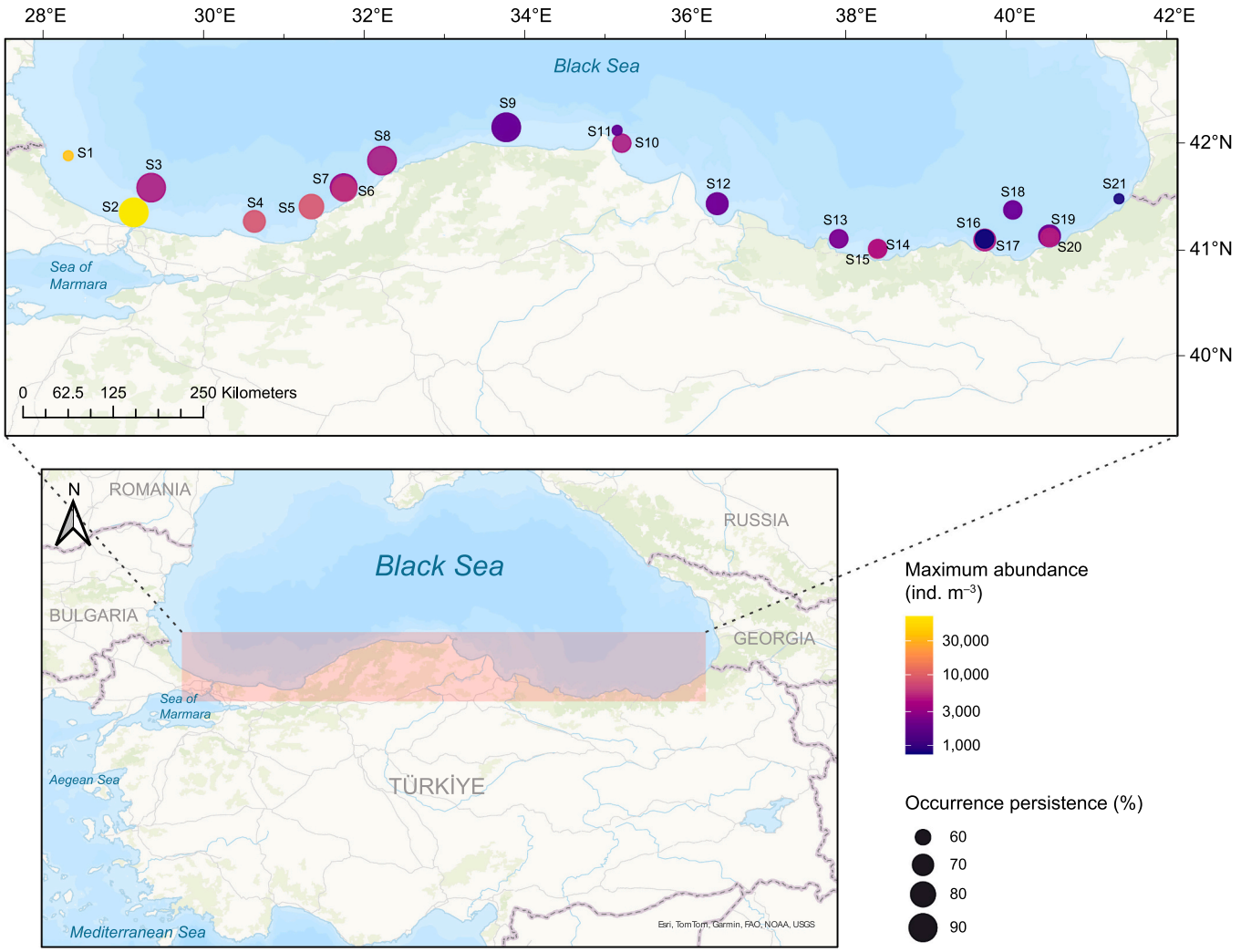

Fig. 1. Fixed monitoring stations and spatial variation in *Noctiluca scintillans* occurrence persistence and maximum abundance along the southern Black Sea coast, 2018–2024. Symbol size represents the percentage of valid abundance observations with detection; symbol colour represents maximum observed abundance (ind. m^{-3}) on a logarithmic scale. Station identifiers, codes, coordinates, depths, and monitoring localities are reported in Table 1. Map created using ArcGIS Pro (Esri, Redlands, CA, USA). Basemap sources: Esri, TomTom, Garmin, FAO, NOAA, and USGS.

complete-case observations.

2.5. Treatment of censored nutrient values

Analytical limits of detection (LODs) were 0.041 µM for NH_4-N, 0.050 µM for $NO_3 + NO_2$-N, 0.020 µM for PO_4-P, and 0.060 µM for silicate. Measurements below the corresponding LOD, including negative values after blank correction, were treated as left-censored. Censored concentrations were replaced by $LOD/\sqrt{2}$ in the primary analyses. Replacement of censored PO_4-P concentrations by LOD/2 was evaluated in a sensitivity analysis.

Below-LOD observations comprised 5 of 292 NH_4-N measurements (1.71%), 1 of 292 $NO_3 + NO_2$-N measurements (0.34%), and 3 of 292 PO_4-P measurements (1.03%). No silicate measurement was below the LOD.

2.6. Statistical analysis

An endpoint-separated modelling framework accommodated the zero-containing, strongly right-skewed abundance distribution by distinguishing occurrence probability, conditional positive abundance, and Bloom90 probability.

Descriptive analyses characterised non-detection prevalence, seasonal contrasts, interannual variability, and station-level recurrence. Abundance was transformed as $\log_{10}(\text{abundance} + 1)$. Seasonal abundance distributions were compared using a Wilcoxon rank-sum test. Seasonal occurrence frequencies were compared using Pearson's chi-squared test with Yates' continuity correction. Interannual and among-station abundance distributions were compared using Kruskal–Wallis tests.

Environmental predictor structure was evaluated using Spearman rank correlations. Absolute correlations ≥ 0.70 were considered indicative of substantial redundancy when variables represented overlapping environmental gradients (Fig. S1); redundant predictors were not included jointly in the endpoint-specific models.

Principal component analysis (PCA) summarised linear covariance gradients among scaled environmental variables, whereas non-metric multidimensional scaling (NMDS) represented multivariate environmental dissimilarity. Environmental vectors were fitted to the NMDS configuration using envfit. Permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) quantified marginal environmental differences associated with season, year, and occurrence status, with permutations restricted within stations. Multivariate dispersion was assessed using permutational analysis of multivariate dispersions (PERMDISP). Distance-based redundancy analysis (dbRDA) evaluated the joint constrained contributions of season, year, occurrence status,

and log10-transformed station depth. These analyses were implemented in vegan (Oksanen et al., 2026), characterised multivariate environmental structure, and did not replace the endpoint-specific regression models.

The primary endpoint-specific models included season, temperature, Chlorophyll-a, PO_4-P, and year where estimable. Chlorophyll-a and PO_4-P were $\log_{10}$-transformed before standardisation. Continuous predictors were converted to z-scores using complete-dataset means and standard deviations (SDs). One temperature SD corresponded to 6.88°C. The complete-dataset mean ± 1 SD extended from 8.83 to 22.6°C, relative to an observed range of 7.51–28.0°C.

Occurrence probability was estimated using a binomial generalised linear model (GLM) with a logit link:

$$\text{logit}\{\Pr(O_i=1)\}=\beta_0+\beta_1\text{Season}_i+\beta_2T_i+\beta_3C_i+\beta_4P_i+\beta_5^{\mathrm{T}}\text{Year}_i, \tag{5}$$

where T_i, C_i, and P_i denote standardised temperature, $\log_{10}$-transformed Chlorophyll-a, and $\log_{10}$-transformed PO_4-P, respectively.

Conditional positive abundance was estimated using a Gaussian linear model (LM) restricted to observations with $A_i > 0$:

$$\log_{10}(A_i)=\beta_0+\beta_1\text{Season}_i+\beta_2T_i+\beta_3C_i+\beta_4P_i+\beta_5^{\mathrm{T}}\text{Year}_i+\varepsilon_i. \tag{6}$$

Bloom90 probability was estimated using a binomial GLM with a logit link:

$$\text{logit}\{\Pr(B_i=1)\}=\beta_0+\beta_1\text{Season}_i+\beta_2T_i+\beta_3C_i+\beta_4P_i. \tag{7}$$

Year was excluded from the primary Bloom90-probability model because only 30 Bloom90 events were available and one year contained no events. Binomial-model coefficients were reported as odds ratios (ORs) with 95% confidence intervals (CIs). Conditional positive-abundance coefficients were reported on the $\log_{10}$-abundance scale with 95% CIs. Winter and 2018 were the reference levels.

Sensitivity to repeated observations within stations was evaluated using station-clustered HC1 covariance estimation for the binary endpoints and generalised estimating equations (GEEs) with an exchangeable within-station correlation structure implemented in geepack (Halekoh et al., 2006). For conditional positive abundance, bias-reduced CR2 covariance estimation with Satterthwaite degrees of freedom was additionally applied using clubSandwich (Pustejovsky, 2026). Primary model estimates were retained when effect direction, magnitude, and inference were concordant with the cluster-aware estimates.

Model adequacy was assessed using simulation-based residual diagnostics implemented in DHARMa (Hartig, 2026) for the binomial models and conventional residual diagnostics for the conditional positive-abundance model. Collinearity was evaluated using variance-inflation diagnostics, with adjusted generalised variance-inflation factors applied to terms with multiple degrees of freedom.

Leave-one-station-out and leave-one-year-out blocked validation assessed spatial and temporal transportability, respectively. Year was omitted from the leave-one-year-out occurrence-probability and conditional positive-abundance models because the held-out categorical level was absent from the corresponding training dataset and could not be estimated. Binary-endpoint performance was quantified using the area under the receiver operating characteristic curve (ROC AUC), calculated using pROC (Robin et al., 2011); the area under the precision–recall curve (PR AUC), calculated using PRROC (Grau et al., 2015); the Brier score; logarithmic loss; calibration intercept; and calibration slope. Conditional positive-abundance performance was quantified using root-mean-square error (RMSE), mean absolute error (MAE), and predictive R^2.

Station-level occurrence persistence and Bloom90 frequency were reported with two-sided Clopper–Pearson exact binomial 95% CIs. Influence analyses excluded 2021, S2/K0, and both jointly while retaining the complete-dataset Bloom90 threshold and complete-dataset predictor-standardisation parameters. A secondary endpoint-definition sensitivity analysis recalculated the Bloom90 threshold after exclusion of 2021. Additional sensitivity analyses excluded the upper 1% of abundance observations and replaced censored PO_4-P concentrations with LOD/2.

Temperature functional form was evaluated using natural splines with three degrees of freedom and generalised additive models (GAMs) fitted by maximum likelihood using mgcv (Wood, 2017). Evaluation considered nested-model tests, Akaike's information criterion (AIC), estimated effective degrees of freedom, smooth-term uncertainty, convergence, and basis-dimension diagnostics. Linear temperature specifications were retained for primary coefficient reporting when nonlinear terms did not improve model fit or could not be estimated adequately.

Analyses were conducted in R version 4.6.1 (R Core Team, 2026) using RStudio version 2026.06.0 + 242 (Posit Team, 2026). Figures were generated using ggplot2 (Wickham, 2016).

3. RESULTS

3.1. Monitoring network and analytical dataset

The monitoring dataset comprised 294 observations, equally distributed between winter and summer before response filtering ($n = 147$ per season). Two observations lacked valid *N. scintillans* abundance measurements, leaving 292 valid abundance observations for occurrence-probability and Bloom90-probability analyses and 211 positive observations for conditional positive-abundance modelling.

Noctiluca scintillans was detected in 211 of 292 valid abundance observations (72.3%). Abundance was strongly right-skewed, with a median of 93 ind. m^{-3}, a mean of 961 ind. m^{-3}, and a maximum of 66,751 ind. m^{-3}. The unrounded complete-dataset Bloom90 and Bloom95 thresholds were 1744.4 and 4322.35 ind. m^{-3}, respectively, identifying 30 Bloom90 events and 15 Bloom95 events (Table 2; Fig. 2).

3.2. Seasonal and interannual variability

Occurrence was higher in winter than in summer (92.5% versus 52.1%; Pearson's chi-squared test: $\chi^2 = 57.474$, $df = 1$, $P < 0.001$; Tables 2 and S1; Fig. 2). Abundance distributions also differed between seasons (Wilcoxon rank-sum test: $W = 15{,}535$, $P < 0.001$), with a higher median in winter than in summer (265 versus 17 ind. m^{-3}).

Despite the higher winter median, mean abundance was lower in winter than in summer (723 versus 1198 ind. m^{-3}) because of infrequent high-abundance summer observations. Maximum abundance reached 10,298 ind. m^{-3} in winter and 66,751 ind. m^{-3} in summer. Winter was therefore characterised by more persistent occurrence and higher median abundance, whereas summer contributed disproportionately to upper-tail abundance.

Abundance distributions differed among years (Kruskal–Wallis test: $\chi^2 = 40.874$, $df = 6$, $P < 0.001$). Occurrence, mean abundance, median abundance, and maximum abundance were highest in 2021 (Table 2; Fig. S2). Interannual heterogeneity persisted after exclusion of 2021 ($\chi^2 = 17.513$, $df = 5$, $P = 0.004$), indicating that 2021 amplified but did not solely account for temporal variation (Table S1).

3.3. Station-level occurrence persistence and upper-tail abundance

Overall abundance distributions did not differ among stations (Kruskal–Wallis test: $\chi^2 = 17.002$, $df = 20$, $P = 0.653$; Table S1). Spatial variation was instead expressed through occurrence persistence and episodic upper-tail abundance.

S2/K0 had 92.9% occurrence persistence and the highest cumulative abundance (99,578 ind. m^{-3}), maximum abundance (66,751 ind. m^{-3}), and Bloom90 frequency (5/14 observations; 35.7%, exact 95% CI

Table 2
Seasonal and interannual occurrence and abundance of *Noctiluca scintillans* along the southern Black Sea coast, 2018–2024.

Group	Level	Valid abundance observations, *n*	Non-detections, *n*	Positive observations, *n*	Occurrence (%)	Mean abundance (ind. m^{-3})	Median abundance (ind. m^{-3})	Maximum abundance (ind. m^{-3})	90th percentile (ind. m^{-3})	95th percentile (ind. m^{-3})
Season	Winter	146	11	135	92.5	723	265	10298	1946	4248
	Summer	146	70	76	52.1	1198	17	66751	1112	4234
Year	2018	41	12	29	70.7	811	169	10298	1027	5243
	2019	42	18	24	57.1	132	24	926	408	597
	2020	42	12	30	71.4	1273	260	8302	3981	5590
	2021	42	1	41	97.6	3779	524	66751	4795	8021
	2022	42	16	26	61.9	192	39	2323	463	701
	2023	42	7	35	83.3	349	170	4954	680	903
	2024	41	15	26	63.4	165	43	2135	350	480

Note. Values are unadjusted observed summaries. Percentiles were calculated separately within each season and year and therefore differ from the complete-dataset Bloom90 and Bloom95 thresholds.

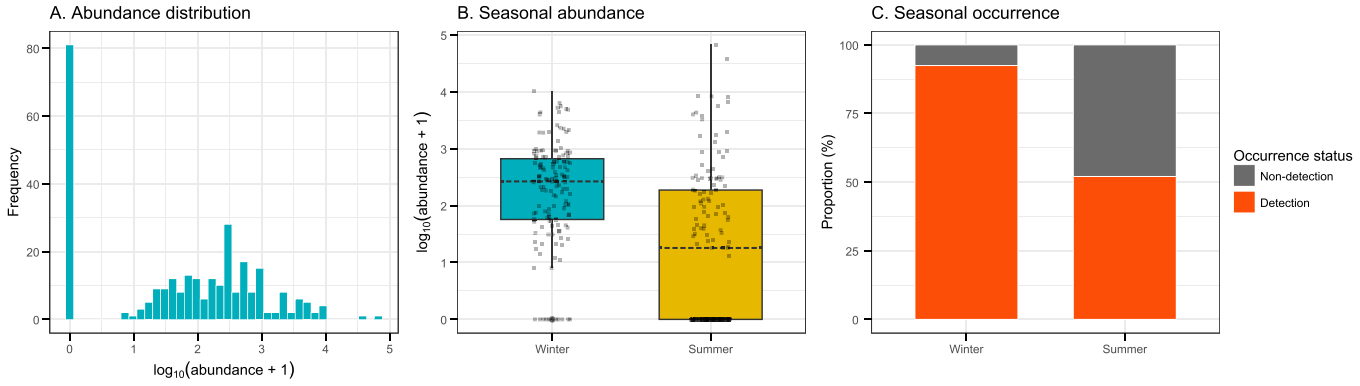

Fig. 2. Distribution and seasonal structure of *Noctiluca scintillans* abundance and occurrence along the southern Black Sea coast, 2018–2024. (A) Frequency distribution of $\log_{10}$(abundance + 1). (B) Seasonal distributions of $\log_{10}$(abundance + 1). Boxes show medians and interquartile ranges; whiskers extend to 1.5 times the interquartile range; points represent individual observations. (C) Seasonal proportions of detection and non-detection observations.

12.8–64.9%). S5/TRKE2, S6/TRK13, and S7/TRK14Y each contained three Bloom90 events (3/14 observations; 21.4%, exact 95% CI 4.7–50.8%). Wide, overlapping CIs precluded precise ranking of station-level Bloom90 frequencies (Table S2; Figs. 1 and 3).

Occurrence persistence exceeded 92% at S8/TRK21, S3/1H15, and S9/TRK23Y, although each station contained only one Bloom90 event. Conversely, S1/TRK2Y had 57.1% occurrence persistence but contained two Bloom90 events and reached a maximum abundance of 37,551 ind. m^{-3}. These contrasts distinguished persistent occurrence from episodic upper-tail abundance.

Bloom90 events were temporally episodic and concentrated at several western stations (Fig. 4). After exclusion of S2/K0, three events remained at each of S5/TRKE2, S6/TRK13, and S7/TRK14Y. S2/K0 therefore strengthened but did not solely account for the concentration of Bloom90 events at stations near the Istanbul Strait outflow and Zonguldak.

3.4. Environmental gradients and occurrence-state separation

Principal component axes 1 and 2 explained 40.3% and 18.1% of environmental variation, respectively, accounting jointly for 58.4% (Table S3; Fig. 5). Non-metric multidimensional scaling provided an adequate two-dimensional representation of environmental dissimilarity (stress = 0.163), and all fitted environmental vectors were significantly related to the NMDS configuration ($P = 0.001$ for each vector).

In marginal PERMANOVA, season explained 15.5% of environmental variation ($R^2 = 0.155$, $F = 66.21$, $P = 0.001$), year explained 11.3% ($R^2 = 0.113$, $F = 8.03$, $P = 0.001$), and occurrence status explained 1.28% ($R^2 = 0.0128$, $F = 5.48$, $P = 0.008$). After inclusion of $\log_{10}$-transformed station depth, depth explained 12.5% of environmental variation ($R^2 = 0.125$, $F = 66.26$, $P = 0.001$), whereas the contribution of occurrence status remained small ($R^2 = 0.0122$, $F = 6.47$, $P = 0.005$; Table S3).

Multivariate dispersion did not differ between non-detection and detection observations (PERMDISP: $F = 0.44$, $P = 0.509$). Seasonal dispersion differed ($F = 7.93$, $P = 0.005$), indicating that seasonal PERMANOVA separation reflected both centroid displacement and dispersion heterogeneity.

Non-detections were concentrated at higher temperatures, whereas positive observations and Bloom90 events occurred predominantly under cooler conditions. Bloom90 events also tended to coincide with higher Chlorophyll-a concentrations, whereas separation among abundance categories was less distinct for PO_4-P (Fig. 6).

Strong Spearman rank correlations occurred between dissolved inorganic nitrogen (DIN) and $NO_3 + NO_2$-N ($\rho = 0.933$), dissolved oxygen and temperature ($\rho = -0.813$), DIN and PO_4-P ($\rho = 0.765$), PO_4-P and silicate ($\rho = 0.761$), and $NO_3 + NO_2$-N and PO_4-P ($\rho = 0.742$; Fig. S1).

3.5. Occurrence probability

Occurrence probability was associated primarily with temperature. Station-clustered and GEE estimates were concordant with the primary model, and the estimated within-station correlation was negligible ($\alpha = -0.016$, SE = 0.033; Table S4).

The occurrence-probability model achieved an in-sample ROC AUC of 0.895 and showed no material residual inadequacy (Table 3; Fig. S3). Leave-one-station-out blocked validation retained good discrimination, with a ROC AUC of 0.860, PR AUC of 0.918, and Brier score of 0.123. The corresponding values under leave-one-year-out blocked validation were 0.780, 0.851, and 0.144, respectively. Calibration slopes of 0.746

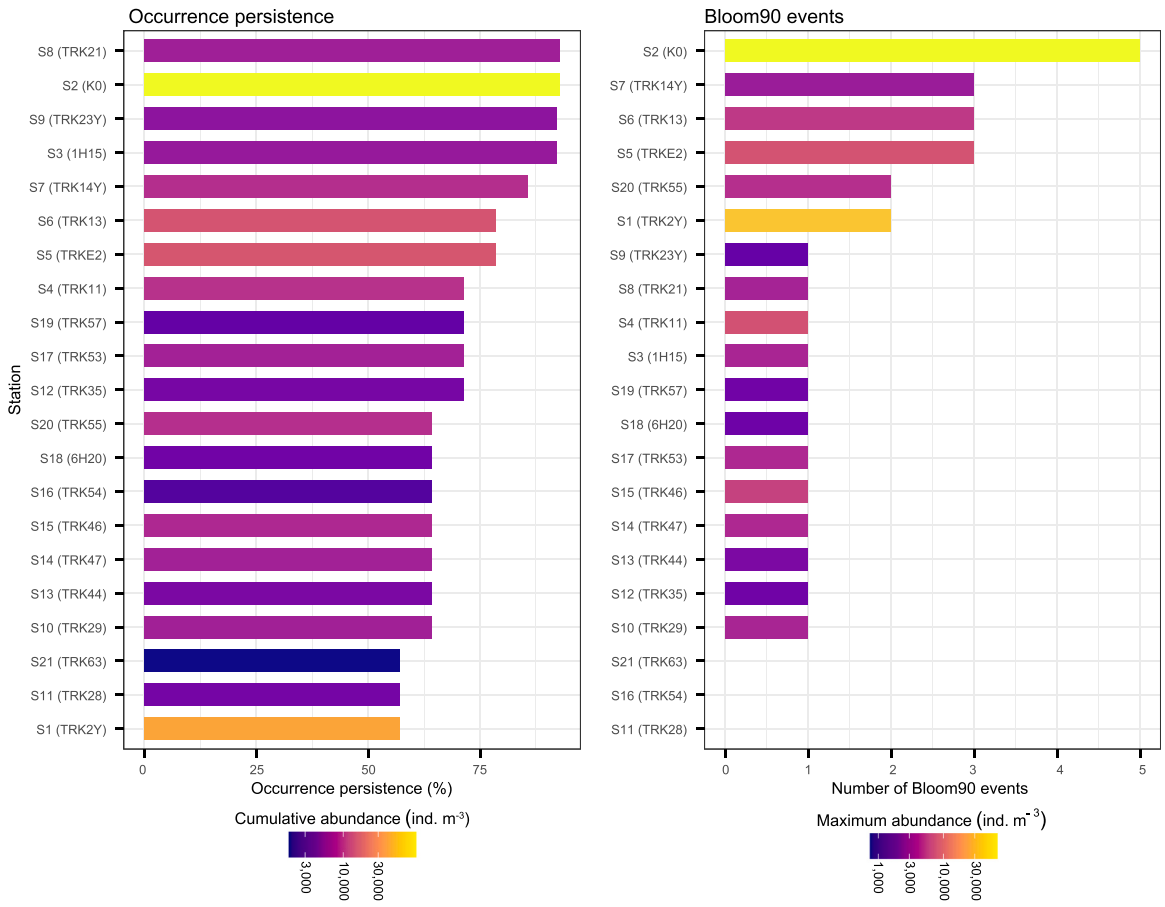

Fig. 3. Station-level occurrence persistence and Bloom90 frequency of *Noctiluca scintillans* along the southern Black Sea coast, 2018–2024. The left panel shows the percentage of valid abundance observations with detection; bar colour represents cumulative abundance (ind. m^{-3}). The right panel shows the number of Bloom90 events; bar colour represents maximum observed abundance (ind. m^{-3}). Both colour scales are logarithmic. Station-level occurrence persistence and Bloom90 frequencies, with two-sided Clopper–Pearson exact 95% confidence intervals, are reported in Table S2.

and 0.760 indicated moderate overfitting under leave-one-station-out and leave-one-year-out blocked validation, respectively (Table S5; Fig. S4).

Each 1-SD increase in temperature was associated with a 91.0% reduction in occurrence odds (OR = 0.090, 95% CI 0.031–0.223, $P < 0.001$; Table 4; Fig. 7). The direction and magnitude of this association were stable across temporal and spatial influence analyses (Fig. 8) and under cluster-aware inference (Table S4).

Relative to 2018, adjusted occurrence odds were lower in 2019 (OR = 0.268, 95% CI 0.073–0.935, $P = 0.042$), 2022 (OR = 0.067, 95% CI 0.013–0.298, $P < 0.001$), and 2024 (OR = 0.207, 95% CI 0.050–0.801, $P = 0.026$). Chlorophyll-a and PO_4-P were not independently associated with occurrence probability (Table 4).

3.6. Conditional positive abundance

The conditional positive-abundance model explained a modest proportion of variation ($R^2 = 0.160$; adjusted $R^2 = 0.118$). Residuals were consistent with normality (Shapiro–Wilk test: $P = 0.895$; Fig. S5), and the estimated within-station correlation was negligible ($\alpha = -0.003$, SE = 0.026; Table S4).

Relative to 2018, conditional positive abundance was lower in 2019 ($\beta = -0.507$, 95% CI −0.889 to −0.125, $P = 0.009$), 2022 ($\beta = -0.461$, 95% CI −0.855 to −0.068, $P = 0.022$), 2023 ($\beta = -0.417$, 95% CI −0.784 to −0.051, $P = 0.026$), and 2024 ($\beta = -0.465$, 95% CI −0.853 to −0.078, $P = 0.019$; Table 4).

Temperature showed a negative but uncertain association with conditional positive abundance ($\beta = -0.238$, 95% CI −0.490–0.014, $P = 0.064$; Table 4; Fig. 9). The coefficient magnitude was retained under CR2 estimation ($\beta = -0.238$, $P = 0.053$), GEE estimation ($\beta = -0.236$, $P = 0.026$), and exclusion of the upper 1% of abundance observations ($\beta = -0.226$, 95% CI −0.465–0.014, $P = 0.065$; Table S4). Chlorophyll-a and PO_4-P were not independently associated with conditional positive abundance.

Leave-one-station-out blocked validation yielded an RMSE of 0.716 $\log_{10}$ units, an MAE of 0.569 $\log_{10}$ units, and a predictive R^2 of 0.073. Leave-one-year-out blocked validation yielded an RMSE of 0.785 $\log_{10}$ units, an MAE of 0.634 $\log_{10}$ units, and a predictive R^2 of −0.117, indicating no temporally transferable predictive performance (Table S5).

3.7. Bloom90 probability

After environmental adjustment, summer observations had 13.7-fold higher Bloom90 odds than winter observations (OR = 13.7, 95% CI 2.93–67.1, $P < 0.001$; Table 4; Fig. 10). This pattern contrasted with higher unadjusted winter occurrence and median abundance, demonstrating endpoint-specific seasonal structure.

Each 1-SD increase in temperature was associated with an 81.8% reduction in Bloom90 odds (OR = 0.182, 95% CI 0.066–0.455, $P < 0.001$). Bloom90 probability was positively associated with $\log_{10}$-transformed Chlorophyll-a (OR = 1.74, 95% CI 1.06–2.92, $P = 0.031$) and negatively associated with $\log_{10}$-transformed PO_4-P (OR = 0.590, 95% CI 0.398–0.878, $P = 0.008$; Table 4). Station-clustered and GEE estimates were concordant with the primary estimates, the estimated within-station correlation was weak ($\alpha = 0.014$, SE = 0.034; Table S4),

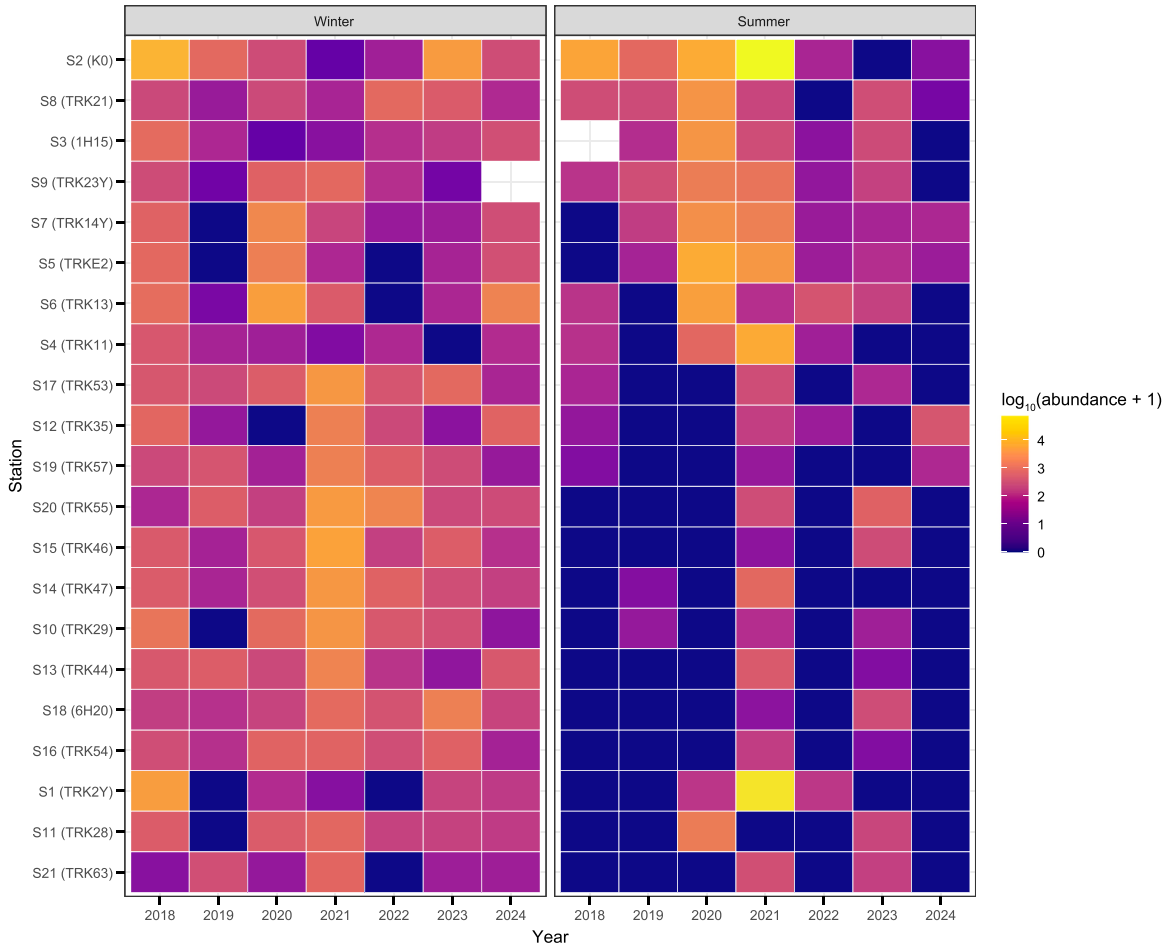

Fig. 4. Station–year–season distribution of *Noctiluca scintillans* abundance along the southern Black Sea coast, 2018–2024. Cell colour represents $\log_{10}$(abundance + 1) for each station–year–season observation. Winter and summer observations are shown in separate panels. Stations are ordered by decreasing occurrence persistence and, where occurrence persistence is equal, by decreasing cumulative abundance.

and simulation-based diagnostics indicated no material residual inadequacy (Fig. S6).

Leave-one-station-out blocked validation yielded a ROC AUC of 0.698, PR AUC of 0.230, Brier score of 0.087, and calibration slope of 0.776. The corresponding values under leave-one-year-out blocked validation were 0.517, 0.099, 0.103, and 0.013, respectively, indicating no meaningful discrimination or calibration under temporal validation (Table S5; Fig. S4). The Bloom90-probability model therefore supported explanatory analysis and monitoring prioritisation within the observed system but not temporally transferable forecasting of Bloom90 events.

3.8. Influence and functional-form analyses

After exclusion of 2021, the occurrence–temperature association remained nearly unchanged (OR = 0.097, 95% CI 0.034–0.240, $P < 0.001$), and interannual variation in conditional positive abundance persisted. Bloom90 associations with summer (OR = 13.2, 95% CI 2.22–84.0, $P = 0.005$) and temperature (OR = 0.142, 95% CI 0.041–0.409, $P < 0.001$) were retained. The complete-data Chlorophyll-a association was not retained (OR = 0.941, 95% CI 0.465–1.99, $P = 0.870$), whereas the negative PO_4-P coefficient remained directionally consistent but was estimated with greater uncertainty (OR = 0.615, 95% profile-likelihood CI 0.390–1.01, Wald $P = 0.040$; Fig. 8).

Recalculation of the Bloom90 threshold after exclusion of 2021 reduced the threshold to 904.1 ind. m^{-3} and identified 25 Bloom90 events. Associations with summer (OR = 5.77, 95% CI 1.18–27.8, $P = 0.027$), temperature (OR = 0.157, 95% CI 0.051–0.417, $P < 0.001$), and PO_4-P (OR = 0.664, 95% CI 0.448–0.999, $P = 0.042$) were retained, whereas Chlorophyll-a was not associated with Bloom90 probability (OR = 0.775, 95% CI 0.433–1.41, $P = 0.393$; Table S4).

Exclusion of S2/K0 retained 278 observations and 25 Bloom90 events. Bloom90 estimates remained comparable with the complete-data estimates for summer (OR = 11.3, 95% CI 2.06–62.7), temperature (OR = 0.192, 95% CI 0.063–0.524), Chlorophyll-a (OR = 1.76, 95% CI 1.04–3.08), and PO_4-P (OR = 0.559, 95% CI 0.366–0.855). Combined exclusion of 2021 and S2/K0 retained the occurrence–temperature association and Bloom90 associations with summer and temperature despite reducing the Bloom90 event count to 13 (Fig. 8). Temperature associations were stable across influence analyses, whereas the complete-data Chlorophyll-a association was sensitive to exclusion of 2021.

Exploratory all-smooth associations for occurrence probability and conditional positive abundance are shown in Figs. S7 and S8, respectively. Formal temperature functional-form comparisons are reported in Table S4 and Fig. S9. Natural-spline and GAM analyses supported an effectively linear occurrence–temperature association. The conditional positive-abundance association was also effectively linear but remained uncertain. Bloom90 probability showed modest departure from linearity: the natural-spline model improved fit relative to the linear model (likelihood-ratio test: $\chi^2 = 6.98$, $df = 2$, $P = 0.031$; AIC = 179 versus 182), and the estimated smooth had an effective degree of freedom of 1.71 ($P = 0.003$). The basis-dimension diagnostic nevertheless indicated

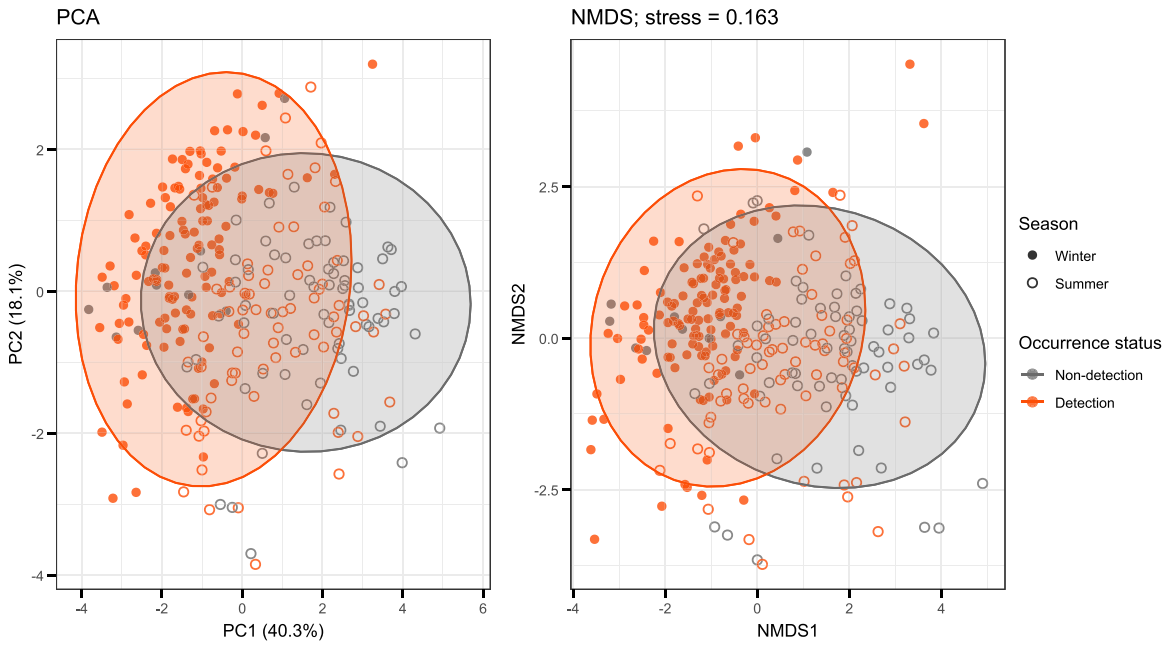

Fig. 5. Ordination of environmental conditions by *Noctiluca scintillans* occurrence status along the southern Black Sea coast, 2018–2024. The left panel shows principal component analysis scores; principal components 1 and 2 explained 40.3% and 18.1% of environmental variation, respectively. The right panel shows the two-dimensional non-metric multidimensional scaling configuration based on Euclidean distances among scaled environmental variables (stress = 0.163). Colours distinguish detection from non-detection observations, symbols distinguish winter from summer observations, and ellipses represent 95% multivariate *t*-distribution concentration regions by occurrence status.

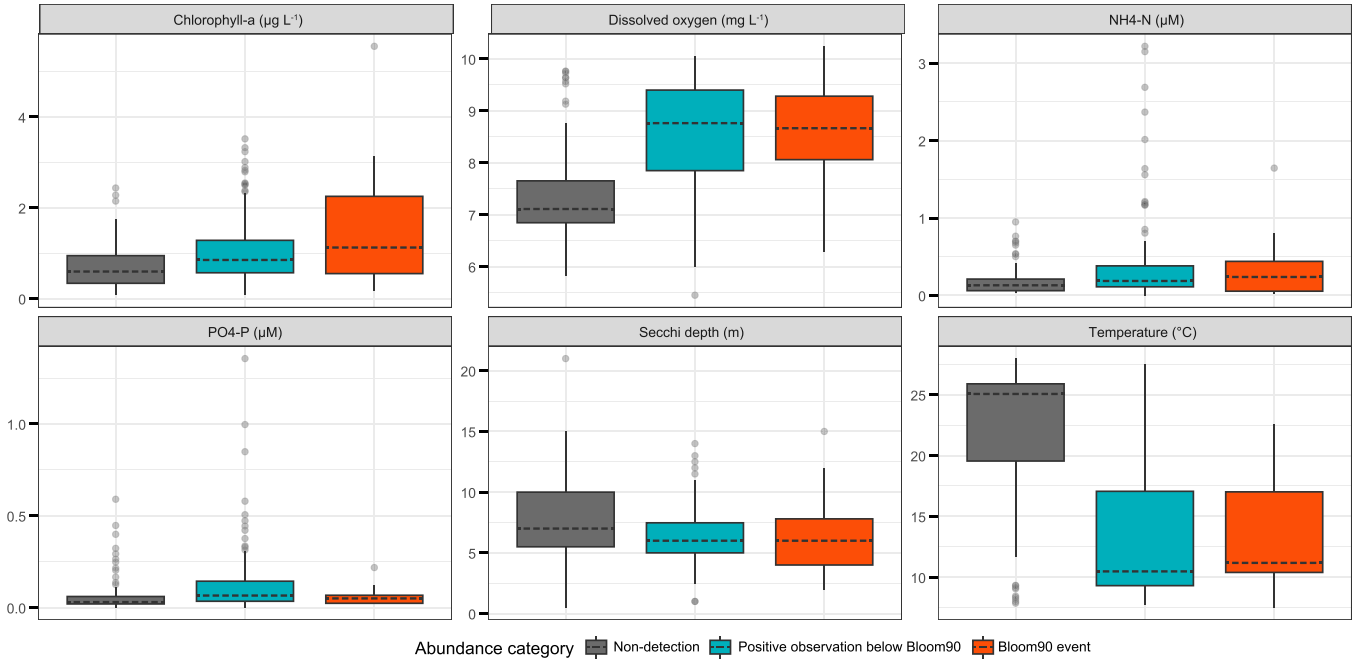

Fig. 6. Environmental conditions across *Noctiluca scintillans* abundance categories along the southern Black Sea coast, 2018–2024. Boxplots show temperature, Chlorophyll-a, phosphate phosphorus (PO_4-P), dissolved oxygen, ammonium nitrogen (NH_4-N), and Secchi depth for non-detections, positive observations below the Bloom90 threshold, and Bloom90 events. Boxes show medians and interquartile ranges; whiskers extend to 1.5 times the interquartile range; points beyond the whiskers represent individual observations.

unresolved residual structure (k-index = 0.81, $P < 0.001$; Table S4; Fig. S9). The linear temperature specification for Bloom90 probability was retained for primary coefficient reporting because the nonlinear analyses did not identify a stable temperature threshold or optimum.

4. DISCUSSION

4.1. Endpoint-specific environmental associations

Occurrence probability, conditional positive abundance, and Bloom90 probability showed distinct environmental and temporal

Table 3
Model specifications, in-sample performance, diagnostic assessment, and cluster-aware sensitivity analyses for *Noctiluca scintillans* occurrence probability, conditional positive abundance, and Bloom90 probability.

Endpoint	Analytical sample	Model specification	In-sample performance	Diagnostic assessment	Cluster-aware sensitivity analysis
Occurrence probability	n= 292; 211 detections	Binomial generalised linear model with logit link: season + year + temperature + $\log_{10}$-transformed Chlorophyll-a + $\log_{10}$-transformed PO_4-P	AIC = 228.36; ROC AUC = 0.895	Simulation-based diagnostics indicated no material departures in residual uniformity or dispersion and no evidence of excessive outlier frequency	Station-clustered and exchangeable-correlation GEE estimates were concordant with the primary model; GEE α = −0.016, SE = 0.033
Conditional positive abundance	n= 211 positive observations	Gaussian linear model: season + year + temperature + $\log_{10}$-transformed Chlorophyll-a + $\log_{10}$-transformed PO_4-P	AIC = 460.72; R^2 = 0.160; Adjusted R^2 = 0.118	Conventional residual diagnostics indicated no material model inadequacy; residuals were consistent with normality (Shapiro–Wilk P = 0.895)	Bias-reduced CR2 covariance estimation and exchangeable-correlation GEE assessed sensitivity to within-station dependence; GEE α = −0.003, SE = 0.026
Bloom90 probability	n= 292; 30 Bloom90 events	Binomial generalised linear model with logit link: season + temperature + $\log_{10}$-transformed Chlorophyll-a + $\log_{10}$-transformed PO_4-P	AIC = 181.89	Simulation-based diagnostics indicated no material departures in residual uniformity or dispersion and no evidence of excessive outlier frequency	Station-clustered and exchangeable-correlation GEE estimates were concordant with the primary model; GEE α = 0.014, SE = 0.034

Note. AIC, Akaike information criterion; CR2, bias-reduced cluster-robust covariance estimator; GEE, generalised estimating equation; PO_4-P, phosphate phosphorus; ROC AUC, area under the receiver operating characteristic curve; SE, standard error; α, estimated exchangeable within-station correlation. Continuous predictors were standardised using complete-dataset means and standard deviations; Chlorophyll-a and PO_4-P were $\log_{10}$-transformed before standardisation. Blocked out-of-sample performance is reported in Table S5.

associations. Occurrence probability represented broad distributional persistence and was associated primarily with temperature, whereas conditional positive abundance varied principally among years. In the complete dataset, Bloom90 probability was associated with summer, lower temperature, higher Chlorophyll-a, and lower PO_4-P. Occurrence and upper-tail abundance therefore represented non-equivalent components of *N. scintillans* dynamics.

Frequent winter occurrence and higher adjusted summer Bloom90 probability further distinguished broad persistence from episodic upper-tail abundance. Bloom90 is a dataset-specific upper-decile abundance endpoint rather than a universal ecological, physiological, toxicological, or harmful-bloom threshold. Poor discrimination and calibration under leave-one-year-out blocked validation restrict its application to explanatory analysis and monitoring prioritisation within the observed system rather than temporally transferable forecasting.

4.2. Temperature associations and regional context

Temperature was the strongest and most consistent environmental correlate of occurrence probability and Bloom90 probability. The occurrence–temperature association remained stable across temporal and spatial influence analyses and under cluster-aware inference. Discrimination remained relatively high under leave-one-station-out blocked validation but declined under leave-one-year-out blocked validation, indicating spatial transportability within the monitored network but reduced temporal transportability.

This pattern accords with high spring abundance and a reported negative abundance–temperature association in the south-eastern Black Sea, as well as with the independent historical event off Rize at 9.9–11.7°C (Kopuz et al., 2014; Şen Ozdemir et al., 2017). Functional-form analyses supported an effectively linear occurrence–temperature association across the observed range. Bloom90 probability showed modest evidence of nonlinearity, but basis-dimension diagnostics indicated unresolved residual structure, and the available data did not identify a stable threshold or optimum. The primary Bloom90 odds ratio therefore represents an average adjusted association across the observed temperature range of 7.51–28.0°C.

These findings do not establish a species-wide thermal preference. *Noctiluca scintillans* is eurythermal, and regional abundance patterns may depend on hydrography, salinity, prey composition, stratification, transport, and bloom phenology (Huang and Qi, 1997; Elbrächter and Qi, 1998; Harrison et al., 2011; Mikaelyan et al., 2014). The estimated associations apply to the monitored southern Black Sea system, sampled seasons, and fixed-station design.

4.3. Seasonal contrast between persistence and upper-tail abundance

High winter occurrence and higher adjusted summer Bloom90 probability indicated distinct seasonal structures for persistence and upper-tail abundance. Cooler winter conditions coincided with widespread occurrence at predominantly low-to-moderate abundance, whereas Bloom90 events occurred disproportionately during summer under a narrower combination of environmental conditions.

Population development followed by physical aggregation provides one plausible explanation. Transport, hydrodynamic retention, surface convergence, and stable meteorological conditions can increase local *N. scintillans* abundance (Huang and Qi, 1997; Dela-Cruz et al., 2003; Kopuz et al., 2014; Mikaelyan et al., 2014). The episodic temporal distribution and concentration of Bloom90 events at several western stations were consistent with this interpretation, but the monitoring data could not partition population growth from advective transport or physical aggregation. Concurrent measurements of biological rates, currents, stratification, wind forcing, and surface convergence would be required to resolve their relative contributions (Dela-Cruz et al., 2003; Kordubel et al., 2024).

4.4. Chlorophyll-a, PO_4-P, and potential trophic time lags

The complete-data association between Chlorophyll-a and Bloom90 probability was not retained after exclusion of 2021 or recalculation of the Bloom90 threshold in the reduced dataset. This association was therefore sensitive to interannual structure. Chlorophyll-a should be interpreted as a bulk phytoplankton/productivity-related proxy rather than a direct measure of edible prey biomass, prey composition, or prey availability (Zhang et al., 2016; Ollevier et al., 2021; Sheng et al., 2022).

The positive complete-data Chlorophyll-a coefficient and negative PO_4-P coefficient were consistent with a productive, phosphate-depleted environmental configuration. The negative PO_4-P association remained directionally consistent after exclusion of 2021 and under the recalculated Bloom90 definition, although uncertainty increased in the sparse-event model. A potential explanation is a trophic time lag in which phytoplankton production reduces dissolved phosphate before subsequent exploitation of particulate food by *N. scintillans* (Kiørboe and Titelman, 1998; Nakamura, 1998; Baliarsingh et al., 2016; Sheng et al., 2022). The independent historical event off Rize supports the biological plausibility of this sequence but does not establish the mechanism underlying the present seven-year associations (Kopuz et al., 2014).

Seasonal covariance, water-mass structure, antecedent nutrient uptake, and nutrient recycling by *N. scintillans* provide alternative explanations (Ara et al., 2013; Drits et al., 2013; Baliarsingh et al., 2016). The absence of measurements of phytoplankton composition, ingestion, food-vacuole contents, and temporal succession precluded attribution of the observed associations to prey-mediated processes. Phosphate-depleted post-production conditions therefore remain a testable hypothesis rather than a demonstrated mechanism.

4.5. Interannual variability and recurrent high-abundance locations

Conditional positive abundance varied principally among years and was explained only modestly by environmental conditions measured during the same monitoring events. Interannual heterogeneity persisted after exclusion of 2021, although that year amplified the complete-data pattern. Year effects represent contrasts among monitoring years rather than a directional temporal trend and may integrate unmeasured hydroclimatic variability, antecedent productivity, grazing, transport, and physical aggregation (Dela-Cruz et al., 2003; Mikaelyan et al., 2014; Kordubel et al., 2024). The negative predictive R^2 under leave-one-year-out blocked validation further indicated that same-event covariates did not support temporally transferable prediction of conditional positive abundance.

S2/K0 had the highest cumulative abundance, maximum abundance, and observed Bloom90 frequency, but its CI overlapped those of other

Table 4
Principal adjusted estimates from models of *Noctiluca scintillans* occurrence probability, conditional positive abundance, and Bloom90 probability.

Endpoint	Predictor contrast	Effect measure	Estimate	95% CI	*P* value
Occurrence probability	Temperature, per 1-SD increase	OR	0.090	0.031–0.223	< 0.001
	2019 versus 2018		0.268	0.073–0.935	0.042
	2022 versus 2018		0.067	0.013–0.298	< 0.001
	2024 versus 2018		0.207	0.050–0.801	0.026
Conditional positive abundance	Temperature, per 1-SD increase	β	−0.238	−0.490–0.014	0.064
	2019 versus 2018		−0.507	−0.889 to −0.125	0.009
	2022 versus 2018		−0.461	−0.855 to −0.068	0.022
	2023 versus 2018		−0.417	−0.784 to −0.051	0.026
	2024 versus 2018		−0.465	−0.853 to −0.078	0.019
Bloom90 probability	Summer versus winter	OR	13.700	2.930–67.100	< 0.001
	Log_{10}-transformed Chlorophyll-a, per 1-SD increase		1.740	1.060–2.920	0.031
	Log_{10}-transformed PO_4-P, per 1-SD increase		0.590	0.398–0.878	0.008
	Temperature, per 1-SD increase		0.182	0.066–0.455	< 0.001

Note. CI, confidence interval; OR, odds ratio; PO_4-P, phosphate phosphorus; SD, standard deviation; β, regression coefficient. Regression coefficients are expressed on the log_{10}-abundance scale. Continuous-predictor estimates represent a 1-SD increase based on complete-dataset standardisation; one temperature SD corresponded to 6.88°C. Winter and 2018 were the reference levels. Estimates were obtained from the primary complete-data models; influence analyses are presented in Fig. 8.

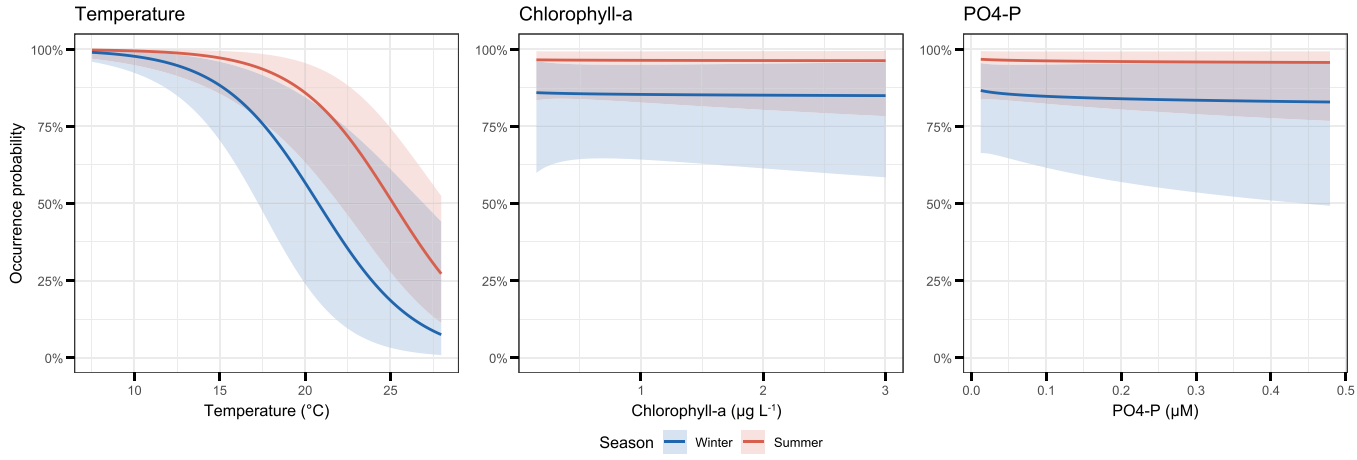

Fig. 7. Adjusted occurrence probability of *Noctiluca scintillans* across temperature, Chlorophyll-a, and phosphate phosphorus (PO_4-P) gradients under winter and summer conditions. Lines and shaded bands represent model-estimated probabilities and 95% confidence intervals from the primary occurrence-probability model. Temperature predictions span the observed range; Chlorophyll-a and PO_4-P predictions span the central 96% of their respective observed distributions. In each panel, non-focal continuous predictors were fixed at their complete-dataset means, and year was fixed at the 2018 reference level.

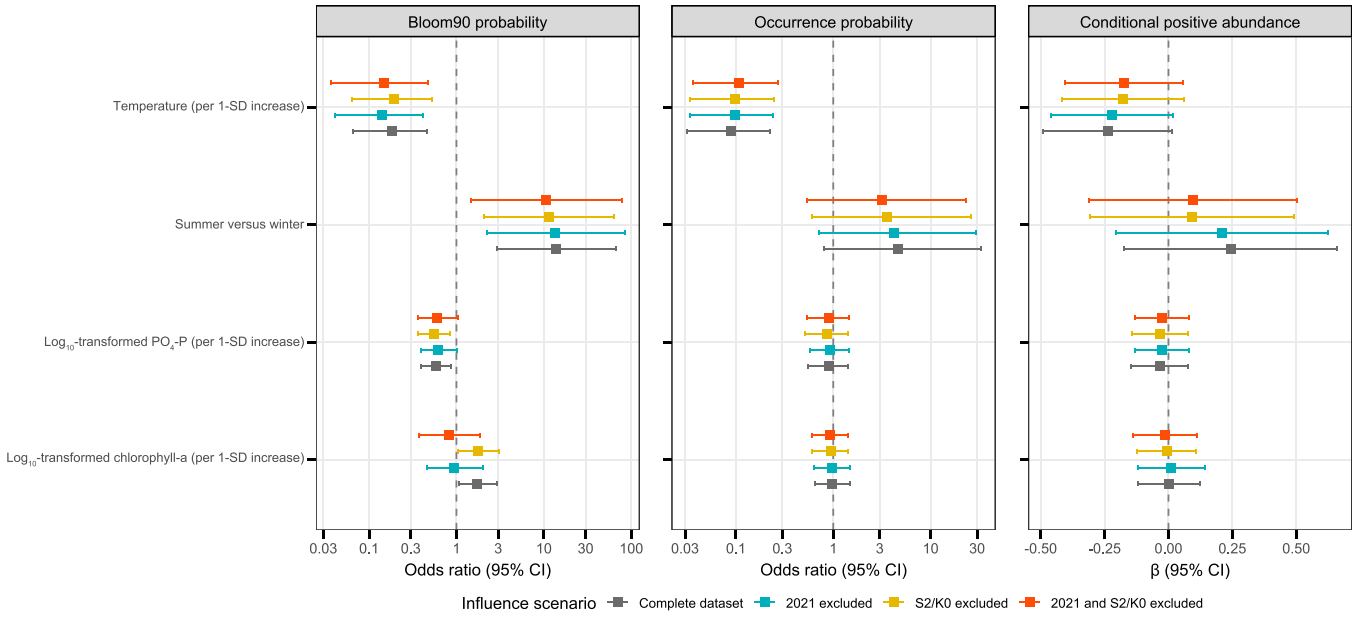

Fig. 8. Influence analyses of adjusted associations with *Noctiluca scintillans* occurrence probability, conditional positive abundance, and Bloom90 probability. Estimates are shown for the complete dataset and after exclusion of 2021, S2/K0, or both 2021 and S2/K0. Binary-endpoint estimates are expressed as odds ratios with profile-likelihood 95% confidence intervals and displayed on logarithmic scales. Conditional positive-abundance estimates are regression coefficients with 95% confidence intervals on the $\log_{10}$-abundance scale. Continuous-predictor effects represent a 1-SD increase based on complete-dataset standardisation. Dashed vertical lines denote an odds ratio of 1 or a regression coefficient of 0.

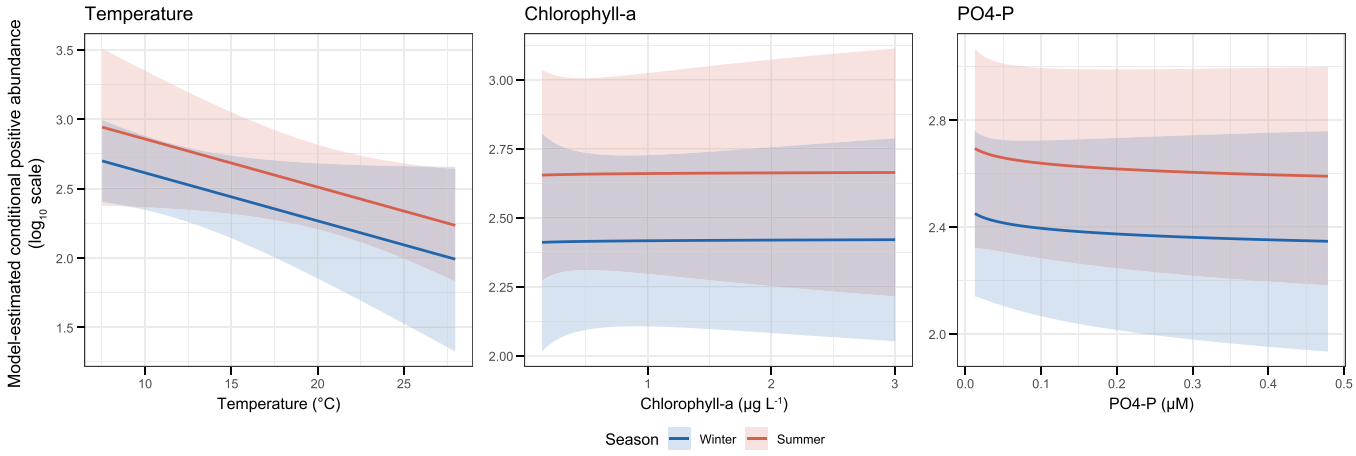

Fig. 9. Adjusted conditional positive abundance of *Noctiluca scintillans* across temperature, Chlorophyll-a, and phosphate phosphorus (PO_4-P) gradients under winter and summer conditions. Lines and shaded bands represent model-estimated conditional positive abundance and 95% confidence intervals on the $\log_{10}$-abundance scale from the primary conditional positive-abundance model fitted to positive observations. Temperature predictions span the observed range; Chlorophyll-a and PO_4-P predictions span the central 96% of their respective observed distributions. In each panel, non-focal continuous predictors were fixed at their complete-dataset means, and year was fixed at the 2018 reference level.

recurrent western stations. Exclusion of S2/K0 did not materially alter the principal Bloom90 estimates, and recurrent Bloom90 events remained at S5/TRKE2, S6/TRK13, and S7/TRK14Y. These stations are therefore best characterised as recurrent high-abundance locations consistent with candidate retention settings rather than statistically distinct locations or hydrodynamically validated retention areas. Convergent circulation, stable surface conditions, and advective accumulation can generate high local *N. scintillans* concentrations, but direct hydrodynamic observations are required to establish retention (Huang and Qi, 1997; Dela-Cruz et al., 2003; Kopuz et al., 2014; Kordubel et al., 2024).

4.6. *Monitoring implications and inferential constraints*

Bloom90 provides an explicit upper-tail state for event-responsive monitoring of potential ecological consequences. Concurrent measurements of harmful algal taxa, food-vacuole contents, and phycotoxins could test ingestion of toxigenic prey and associated phycotoxin-vector potential (Escalera et al., 2007). Measurements of microbial assemblages, dissolved oxygen, ammonium, cell-size structure, and bloom-decay chemistry could directly evaluate microbial and biogeochemical changes accompanying upper-tail abundance (Baliarsingh et al., 2016; Wang et al., 2025; Sathishkumar et al., 2026). Concurrent hydrodynamic observations would additionally be required to partition population development from advective transport and physical aggregation (Dela-Cruz et al., 2003; Kordubel et al., 2024). Bloom90 status

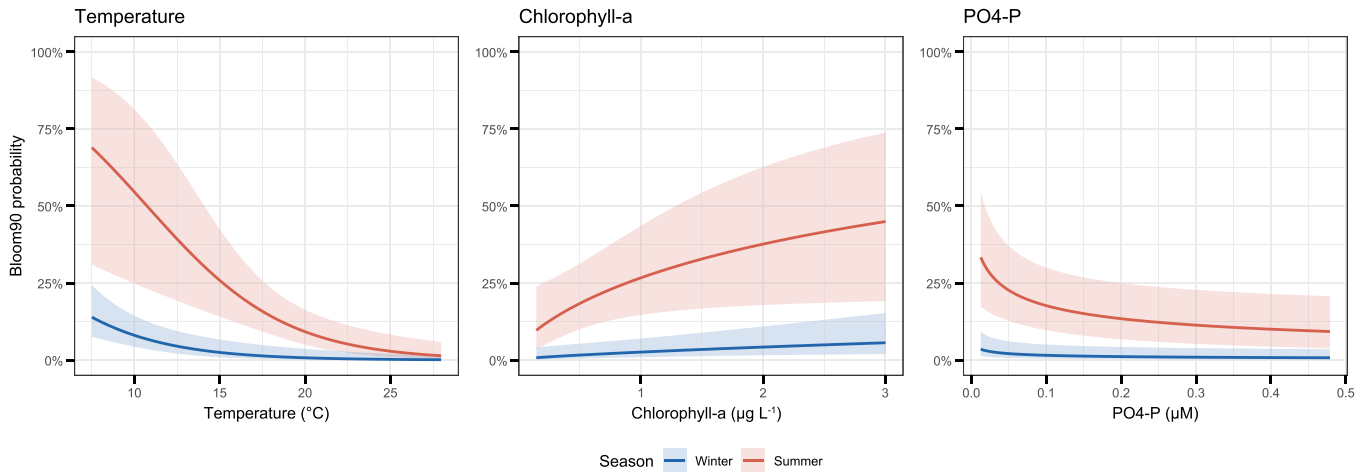

Fig. 10. Adjusted Bloom90 probability of *Noctiluca scintillans* across temperature, Chlorophyll-a, and phosphate phosphorus (PO_4-P) gradients under winter and summer conditions. Lines and shaded bands represent model-estimated probabilities and 95% confidence intervals from the primary Bloom90-probability model. Temperature predictions span the observed range; Chlorophyll-a and PO_4-P predictions span the central 96% of their respective observed distributions. In each panel, non-focal continuous predictors were fixed at their complete-dataset means.

alone does not demonstrate prey-mediated phycotoxin transfer, microbial restructuring, oxygen stress, nutrient recycling, or food-web effects.

Fixed-station sampling resolved broad occurrence patterns and recurrent high-abundance locations but may have missed bloom initiation, short-lived surface aggregations, and spring transitions (Kordubel et al., 2024). Event-responsive monitoring supported by automated imaging, underwater imaging systems, and satellite observations could improve temporal and spatial resolution (Kordubel et al., 2024; Natarajan et al., 2024). Such monitoring should remain hypothesis-driven because the Bloom90-probability model showed no meaningful discrimination or calibration under leave-one-year-out blocked validation.

Mechanistic inference was constrained by the absence of spring sampling, direct prey measurements, toxigenic-taxon data, hydrodynamic observations, microbial profiles, and bloom-process measurements. Nutrient collinearity further restricted simultaneous evaluation of overlapping biogeochemical pathways. Moreover, 23 of the 30 Bloom90 events occurred during 2020–2021, limiting temporal transportability and increasing the sensitivity of some complete-data coefficients to interannual structure. These constraints limit mechanistic and prospective inference but do not negate the endpoint-specific associations estimated within the monitoring dataset.

4.7. Synthesis

The southern Black Sea observations distinguished cool-condition occurrence, interannually variable conditional positive abundance, and episodic upper-tail abundance. The occurrence–temperature association was the most stable environmental signal. Bloom90 associations with summer and temperature were retained across temporal and spatial influence analyses, whereas the complete-data Chlorophyll-a association was sensitive to exclusion of 2021. Recurrent Bloom90 events at several stations near the Istanbul Strait outflow and Zonguldak supported spatially targeted monitoring but not definitive station ranking or hydrodynamic attribution.

The endpoint-separated modelling framework provides an explanatory basis for monitoring prioritisation. Bloom90 defines a dataset-specific upper-tail abundance state for direct investigation of trophic, microbial, biogeochemical, and physical processes; it does not establish harmfulness, prey-mediated phycotoxin transfer, microbial restructuring, hydrodynamic retention, or future-year Bloom90 events.

5. Conclusions

Occurrence probability, conditional positive abundance, and Bloom90 probability represented distinct components of *N. scintillans* dynamics along the southern Black Sea coast. Occurrence probability was strongly and negatively associated with temperature, with good discrimination under leave-one-station-out blocked validation but lower discrimination under leave-one-year-out blocked validation. Conditional positive abundance varied principally among years and showed weak performance under leave-one-station-out blocked validation and no temporally transferable predictive performance under leave-one-year-out blocked validation. In the complete dataset, Bloom90 probability was associated with summer, lower temperature, higher Chlorophyll-a, and lower PO_4-P but showed no meaningful discrimination or calibration under leave-one-year-out blocked validation.

The occurrence–temperature association and Bloom90 associations with summer and temperature remained stable across influence analyses and cluster-aware inference. By contrast, the complete-data association between Chlorophyll-a and Bloom90 probability was not retained after exclusion of 2021 or recalculation of the Bloom90 threshold in the reduced dataset. S2/K0 had the highest observed Bloom90 frequency, but overlapping CIs and recurrent events at other western stations precluded precise spatial ranking.

Bloom90 is a dataset-specific upper-decile abundance endpoint, not a universal harmful-bloom threshold or temporally validated forecasting endpoint. Its principal value lies in defining an explicit upper-tail state for explanatory analysis and event-responsive monitoring. Direct measurements of harmful algal taxa, phycotoxins, microbial assemblages, dissolved oxygen, ammonium, food-vacuole contents, bloom-decay chemistry, and physical aggregation are required to determine whether upper-tail abundance coincides with prey-mediated phycotoxin transfer, microbial restructuring, nutrient recycling, oxygen stress, or broader food-web effects.

Funding

This work was supported by the Integrated Marine Pollution Monitoring Programme 2018–2024, carried out by the Department of Laboratory, Measurement and Monitoring of the General Directorate of Environmental Impact Assessment (EIA), Permit and Inspection, Ministry of Environment, Urbanization and Climate Change, and coordinated by the TÜBİTAK Marmara Research Center (TÜBİTAK MAM).

CRediT authorship contribution statement

İlknur Yıldız: Conceptualization, Data curation, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing. **Ömerhan Dürrani:** Formal analysis, Methodology, Validation, Visualisation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The General Directorate of Environmental Impact Assessment (EIA), Permit and Inspection of the Ministry of Environment, Urbanization and Climate Change of the Republic of Türkiye authorised the use of the DEN-İZ data for publication (Ref. No. E-41234558–604.01.99-16325875). The authors thank two anonymous reviewers for their detailed and constructive comments, which improved the statistical evaluation, interpretation, and presentation of the study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.rsma.2026.105514.

Data availability

Data will be made available on request.

References

Alkan, A., Serdar, S., Fidan, D., Akbaş, U., Zengin, B., Kiliç, M.B., 2022. Spatial, temporal, and vertical variability of nutrients in the Southeastern Black Sea. Chemosphere 302, 134809. https://doi.org/10.1016/j.chemosphere.2022.134809.

Alkan, A., Zengin, B., Serdar, S., Oğuz, T., 2013. Long-Term (2001-2011) Temperature, Salinity and Chlorophyll-a Variations at a Southeastern Coastal Site of the Black Sea. Turk. J. Fish. Aquat. Sci. 13. https://doi.org/10.4194/1303-2712-v13_1_08.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x.

Ara, K., Nakamura, S., Takahashi, R., Shiomoto, A., Hiromi, J., 2013. Seasonal variability of the red tide-forming heterotrophic dinoflagellate *Noctiluca scintillans* in the neritic area of Sagami Bay, Japan: its role in the nutrient-environment and aquatic ecosystem. Plankton Benthos Res. 8, 9–30. https://doi.org/10.3800/pbr.8.9.

Baliarsingh, S.K., Lotliker, A.A., Trainer, V.L., Wells, M.L., Parida, C., Sahu, B.K., Srichandan, S., Sahoo, S., Sahu, K.C., Kumar, T.S., 2016. Environmental dynamics of red *Noctiluca scintillans* bloom in tropical coastal waters. Mar. Pollut. Bull. 111, 277–286. https://doi.org/10.1016/j.marpolbul.2016.06.103.

Dela-Cruz, J., Middleton, J.H., Suthers, I.M., 2003. Population growth and transport of the red tide dinoflagellate, *Noctiluca scintillans*, in the coastal waters off Sydney Australia, using cell diameter as a tracer. Limnol. Oceanogr. 48, 656–674. https://doi.org/10.4319/lo.2003.48.2.0656.

Drits, A.V., Nikishina, A.B., Sergeeva, V.M., Solov'ev, K.A., 2013. Feeding, respiration, and excretion of the Black Sea *Noctiluca scintillans* MacCartney in summer. Oceanology 53, 442–450. https://doi.org/10.1134/s0001437013040036.

Elbrächter, M., Qi, Z., 1998. In: Anderson, D.M., Cembella, A.D., Hallegraeff, G.M. (Eds.), Aspects of Noctiluca (Dinophyceae) population dynamics, Physiological ecology of harmful algal blooms. Springer, Berlin, Heidelberg.

Escalera, L., Pazos, Y., Moroño, Á., Reguera, B., 2007. *Noctiluca scintillans* may act as a vector of toxigenic microalgae. Harmful Algae 6, 317–320. https://doi.org/10.1016/j.hal.2006.04.006.

Grau, J., Grosse, I., Keilwagen, J., 2015. PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics 31, 2595–2597. https://doi.org/10.1093/bioinformatics/btv153.

Halekoh, U., Højsgaard, S., Yan, J., 2006. The R Package geepack for Generalized Estimating Equations. J. Stat. Softw. 15. https://doi.org/10.18637/jss.v015.i02.

Harrison, P.J., Furuya, K., Glibert, P.M., Xu, J., Liu, H.B., Yin, K., Lee, J.H.W., Anderson, D.M., Gowen, R., Al-Azri, A.R., Ho, A.Y.T., 2011. Geographical distribution of red and green *Noctiluca scintillans*. Chin. J. Oceanol. Limnol. 29, 807–831. https://doi.org/10.1007/s00343-011-0510-z.

Hartig, F., 2026. DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models (R package version 0.5.0). CRAN. https://doi.org/10.32614/CRAN.package.DHARMa.

Huang, C., Qi, Y., 1997. The abundance cycle and influence factors on red tide phenomena of *Noctiluca scintillans* (Dinophyceae) in Dapeng Bay, the South China Sea. J. Plankton Res. 19, 303–318. https://doi.org/10.1093/plankt/19.3.303.

Kiørboe, T., Titelman, J., 1998. Feeding, prey selection and prey encounter mechanisms in the heterotrophic dinoflagellate *Noctiluca scintillans*. J. Plankton Res. 20, 1615–1636. https://doi.org/10.1093/plankt/20.8.1615.

Kopuz, U., Feyzioglu, A.M., Valente, A., 2014. An Unusual Red-Tide Event of *Noctiluca scintillans* (Macartney) in the Southeastern Black Sea. Turk. J. Fish. Aquat. Sci. 14, 261–268. https://doi.org/10.4194/1303-2712-v14_1_28.

Kordubel, K., Baschek, B., Hieronymi, M., Voynova, Y.G., Möller, K.O., Dolan, J., 2024. Improving the sampling of red *Noctiluca scintillans* to understand its impact on coastal ecosystem dynamics. J. Plankton Res. 46, 251–271. https://doi.org/10.1093/plankt/fbae010.

Kovalev, A.V., Piontkovski, S.A., 1998. Interannual changes in the biomass of the Black Sea gelatinous zooplankton. J. Plankton Res. 20, 1377–1385. https://doi.org/10.1093/plankt/20.7.1377.

Mikaelyan, A.S., Malej, A., Shiganova, T.A., Turk, V., Sivkovitch, A.E., Musaeva, E.I., Kogovšek, T., Lukasheva, T.A., 2014. Populations of the red tide forming dinoflagellate *Noctiluca scintillans* (Macartney): A comparison between the Black Sea and the northern Adriatic Sea. Harmful Algae 33, 29–40. https://doi.org/10.1016/j.hal.2014.01.004.

Miyaguchi, H., Fujiki, T., Kikuchi, T., Kuwahara, V.S., Toda, T., 2006. Relationship between the bloom of *Noctiluca scintillans* and environmental factors in the coastal waters of Sagami Bay, Japan. J. Plankton Res. 28, 313–324. https://doi.org/10.1093/plankt/fbi127.

Nakamura, Y., 1998. Biomass, feeding and production of *Noctiluca scintillans* in the Seto Inland Sea, Japan. J. Plankton Res. 20, 2213–2222. https://doi.org/10.1093/plankt/20.11.2213.

Natarajan, L., Chandrasekaran, M., Vajravelu, M., Shah, C., Sivadas, S.K., Ramu, K., Murthy, M.V.R., 2024. High Resolution Sentinel-2 and Sentinel-3 Satellite Imagery in Monitoring Green *Noctiluca scintillans* Blooms in Complex Coastal Waters: A Case Study in Gulf of Mannar. J. Indian. Soc. Remote. Sens. 53, 791–802. https://doi.org/10.1007/s12524-024-02032-8.

Nikishina, A.B., Drits, A.V., Vasilyeva, Y.V., Timonin, A.G., Solovyev, K.A., Ratkova, T. N., Sergeeva, V.M., 2011. Role of the *Noctiluca scintillans* population in the trophic dynamics of the Black Sea plankton over the spring period. Oceanology 51, 1029–1039. https://doi.org/10.1134/s0001437011060129.

Oguz, T., Velikova, V., 2010. Abrupt transition of the northwestern Black Sea shelf ecosystem from a eutrophic to an alternative pristine state. Mar. Ecol. Progress. Ser. 405, 231–242. https://doi.org/10.3354/meps08538.

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., De Caceres, M., Durand, S., Evangelista, H.B.A., Hannigan, G., Hill, M.O., Lahti, L., Martino, C., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C.J.F., Weedon, J., 2026. The vegan package. Community Ecol. Package 10, 719. https://doi.org/10.32614/CRAN.package.vegan.

Ollevier, A., Mortelmans, J., Aubert, A., Deneudt, K., Vandegehuchte, M.B., 2021. *Noctiluca scintillans*: Dynamics, Size Measurements and Relationships With Small Soft-Bodied Plankton in the Belgian Part of the North Sea. Front. Mar. Sci. 8. https://doi.org/10.3389/fmars.2021.777999.

Porumb, F., 1992. On the development of *Noctiluca scintillans* under eutrophication of Romanian Black Sea waters. In: Vollenweider, R.A., Marchetti, R., Viviani, R. (Eds.), Marine Coastal Eutrophication. Elsevier, Amsterdam, pp. 907–920.

Posit Team. (2026). RStudio: Integrated development environment for R (Version 2026.06.0+242) [Computer software]. Posit Software, PBC. https://posit.co/.

Pustejovsky, J.E., 2026. clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections (R package version 0.7.0). CRAN. https://doi.org/10.32614/CRAN.package.clubSandwich.

R Core Team, 2026. R: A language and environment for statistical computing. R. Found. Stat. Comput. https://www.R-project.org/.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Muller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinforma. 12, 77. https://doi.org/10.1186/1471-2105-12-77.

Sathishkumar, R.S., Arayillath, A., Ramakrishnan, S.G., Chandrasekaran, M., Shah, C., Natarajan, L., Sivadas, S.K., Ramu, K., 2026. Impact of *Noctiluca scintillans* bloom and decay phases on coastal water quality and zooplankton dynamics in Indian coastal waters. Front. Mar. Sci. 13. https://doi.org/10.3389/fmars.2026.1706039.

Şen Ozdemir, N., Caf, F., Feyzioglu, A.M., Yildiz, I., 2017. Biochemical Content (Fatty Acids, Sterols, Lipophilic Vitamins, Total Protein, MDA, GSH, GSSG) of *Noctiluca scintillans* in the Southeastern Black Sea. Turk. J. Fish. Aquat. Sci. 17, 301–311. https://doi.org/10.4194/1303-2712-v17_2_09.

Sheng, L., Jiang, Z., Sun, Z., Zhu, Y., Zhai, H., Ding, L., Tong, M., Chen, J., Chen, Q., Zeng, J., 2022. *Noctiluca scintillans* Distribution Largely Regulated by Phytoplankton Biomass in the East China Sea and Southern Yellow Sea. Front. Mar. Sci. 9. https://doi.org/10.3389/fmars.2022.899334.

Shiganova, T., Musaeva, E., Arashkevich, E., Kamburska, L., Stefanova, K., Mihneva, V., Polishchuk, L., Timofte, F., Ustun, F., Oguz, T., Khalvashi, M., Tarkan, A.N., 2008. The state of zooplankton. State of the environment of the Black Sea (2001–2006/7). Commission on the Protection of the Black Sea Against Pollution, pp. 201–246.

Wang, Y., Zhou, M., Yue, X., Chen, Y., Su, D., Liu, Z., 2025. *Noctiluca scintillans* Bloom Reshapes Microbial Community Structure, Interaction Networks, and Metabolism Patterns in Qinhuangdao Coastal Waters, China. Microorganisms 13, 1959. https://doi.org/10.3390/microorganisms13081959.

Wickham, H. (2016). Elegant graphics for data analysis. https://doi.org/10.1007/978-3-319-24277-4.
Wood, S.N. (2017). *Generalized Additive Models*.
Yücel, M., Bitir Soylu, B., Ermiş, E., Alımlı, N., Gülmez, Z.A., 2026. The Black Sea as an Extreme Habitat of Earth's Ocean: Biogeochemical Functioning, Unique Ecosystems, and Astrobiological Relevance. Annu. Rev. Earth Planet. Sci. 54, 411–436. https://doi.org/10.1146/annurev-earth-032524-011942.
Zhang, S., Liu, H., Guo, C., Harrison, P.J., 2016. Differential feeding and growth of *Noctiluca scintillans* on monospecific and mixed diets. Mar. Ecol. Progress. Ser. 549, 27–40.