Harmful Algae 121 (2023) 102363

Contents lists available at ScienceDirect

Harmful Algae

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

Quantifying Spatio-temporal risk of Harmful Algal Blooms and their impacts on bivalve shellfish mariculture using a data-driven modelling approach

Oliver Stoner [a], Theo Economou [b], Ricardo Torres [c], Ian Ashton [d], A. Ross Brown [e,*]

[a] School of Mathematics and Statistics, University of Glasgow, UK
[b] Climate and Atmospheric Research Centre, The Cyprus Institute, Nicosia, Cyprus
[c] Plymouth Marine Laboratory, Prospect Place, Plymouth, UK
[d] Department of Engineering, University of Exeter, UK
[e] Department of Biosciences, University of Exeter, UK

ARTICLE INFO

Editor: Dr. C. Gobler

Keywords:
Dinophysis toxins
HAB risk
Official Control monitoring
Marine spatial planning
Statistical modelling

ABSTRACT

Harmful algal blooms (HABs) intoxicate and asphyxiate marine life, causing devastating environmental and socio-economic impacts, costing at least $8bn/yr globally. Accumulation of phycotoxins from HAB phytoplankton in filter-feeding shellfish can poison human consumers, prompting harvesting closures at shellfish production sites. To quantify long-term intoxication risk from *Dinophysis* HAB species, we used historical HAB monitoring data (2009–2020) to develop a new modelling approach to predict *Dinophysis* toxin concentrations in a range of bivalve shellfish species at shellfish sites in Western Scotland, South-West England and Northern France. A spatiotemporal statistical modelling framework was developed within the Generalized Additive Model (GAM) framework to quantify long-term HAB risks for different bivalve shellfish species across each region, capturing seasonal variations, and spatiotemporal interactions. In all regions spatial functions were most important for predicting seasonal HAB risk, offering the potential to inform optimal siting of new shellfish operations and safe harvesting periods for businesses. A 10-fold cross-validation experiment was carried out for each region, to test the models' ability to predict toxin risk at harvesting locations for which data were withheld from the model. Performance was assessed by comparing ranked predicted and observed mean toxin levels at each site within each region: the correlation of ranks was 0.78 for Northern France, 0.64 for Western Scotland, and 0.34 for South-West England, indicating our approach has promise for predicting unknown HAB risk, depending on the region and suitability of training data.

1. Introduction

Aquaculture, including the farming of finfish, shellfish and seaweeds will play a key role in future food security. Global aquaculture production has recently overtaken capture fisheries by producing 82 million tonnes of food fish per year globally (worth US$232 billion), with projections rising to 109 million tonnes by 2030 (FAO, 2020). Marine aquaculture (mariculture) shows enormous potential for sustainable food production (Costello et al., 2020). Expanding the farming of marine bivalve shellfish (currently amounting to 17.3 million tonnes, worth over $24 billion per year globally) is particularly attractive, since these filter feeders derive their food from freely available marine planktonic microalgae; shellfish also perform a wide range of ecosystem services, including nutrient regulation and carbon capture (FAO, 2020; Van der Schatte Olivier et al., 2020).

However, future expansion of bivalve mariculture depends on several socio-economic and environmental constraints relating to competing demands from other marine economic activities (Daniels et al., 2020), environmental carrying capacities (Ferreira et al., 2009) and environmental hazards, including adverse climatic conditions, pollution, and Harmful Algal Blooms (HABs) (Brown et al., 2020). The increasingly frequent and widespread detection of HABs, proliferations of harmful planktonic algae causing intoxication and/or asphyxiation of marine life, is a major constraint on bivalve shellfish mariculture in NW European shelf seas and other HAB hotspots around the globe (Glibert et al., 2014; Weisberg et al., 2019; Trainer et al., 2020; Wells et al., 2020;

* Corresponding author at: Lab 201, Geoffrey Pope Building, Biosciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK.
E-mail address: ross.brown@exeter.ac.uk (A.R. Brown).

https://doi.org/10.1016/j.hal.2022.102363
Received 13 June 2022; Received in revised form 29 November 2022; Accepted 29 November 2022
1568-9883/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Belin et al., 2021; Hallegraeff et al., 2021). Global annual economic impacts of HABs exceed $8 billion, approximately half due to enforced seafood harvesting bans or product recalls and half due to unavoided HAB toxin-related human health costs (GESAMP, 2001; Berdalet et al., 2016). Whilst there is considerable uncertainty in quantifying the global economic costs of HABs (Trainer et al., 2020), the scale of the problem is highlighted by the cost of HAB monitoring, which for example amounted to $6.9 million in Chile alone in 2019 (Mardones et al., 2020). Monitoring costs are only a small fraction of the economic impacts of HABs on fisheries and aquaculture and human health.

In NW Europe, dinoflagellate HAB species from the genus *Dinophysis* (including *Dinophysis acuminata* and *Dinophysis acuta*) are particularly problematic, since low biomass blooms of these species >100 cells/L are sufficient to intoxicate bivalves and cause diarrhetic shellfish poisoning in human consumers (Manfrin et al., 2012; Reguera et al., 2014). These low biomass blooms are undetectable by satellite surveillance and, as with several other dinoflagellate HAB species, they are likely to increase in prevalence with global warming, seasonal thermal stratification, and stabilisation of the water column (Glibert et al., 2014; Wells et al., 2015; Gobler et al., 2017; Trainer et al., 2020).

There is an urgent need for new tools to better quantify long-term changes in HAB risk for shellfish cultivation and harvesting, to inform mitigation strategies, including optimal siting of new shellfish production sites. Current assessment and management of HAB risk, including for *Dinophysis* spp., relies on regular weekly or bi-weekly in situ sampling and subsequent chemical analysis of phycotoxins in shellfish, alongside microscopic quantification of HAB species abundance at routine monitoring points across regulatory networks e.g. in the UK and EU (EC 2004a, b, c). Predicting HAB events and geographical hotspots is highly complex, due to multiple environmental drivers varying temporally from seconds to decades; and spatially, from the microscopic scale (mm) to the mesoscale (100 km). These include physical factors driving water stratification and nutrient depletion, as well as mixing and nutrient replenishment (Anderson et al., 2011; Anderson et al., 2019). Ecological interactions between HAB species, other planktonic organisms and their physico-chemical environments can also be highly influential, notably resource competition, predation, and parasitism – which can invoke biochemical (allelopathic) defence mechanisms involving phycotoxins (Anderson et al., 2011; Wells et al., 2020).

Existing HAB models account for spatial and temporal variation in HAB risk either: explicitly in the case of dispersion models (Davidson et al. 2016) and mechanistic models incorporating key biogeochemical and ecological (life-history) processes (Gillibrand et al., 2016); or implicitly in the case of data models (i.e. statistical or machine learning models) that incorporate short-term and long-term trends (Karasiewicz et al., 2020; Cruz et al., 2021; Fernandes-Salvador, 2021). Data models are simpler and less subject to structural errors, at the expense of overlooking the afore mentioned processes. Moreover, coupled with probabilistic elements, data models can quantify the uncertainty associated with any estimates or predictions. Recent approaches have demonstrated skill in making reliable short-term predictions, notably via flexible smoothing functions (Schmidt et al., 2018) or threshold functions accounting for sudden shifts, e.g. when sea surface temperature falls below tolerable or optimal physiological limits for a HAB species (Taranu et al., 2017). Support vector machines, random forests, probabilistic graphical models, and artificial neural networks have also been shown to be capable of modelling highly dynamic (non-linear) and multi-source physico-chemical and biological data underlying HAB development and decay (Lee et al., 2003; Grasso et al., 2019; Cruz et al., 2021). Long-term recurring HAB hotspots, including for low biomass HAB species such as *Dinophysis* spp., have also been predicted based on physical processes statistically associated with HABs (e.g. through general additive modelling) (Díaz *et al.*, 2019; Fernandes-Salvador *et al.*, 2021).

Here we develop a novel data modelling approach to quantify long-term, recurring risk of shellfish intoxication from *Dinophysis* spp., using 10–20 years of 'Official Control' monitoring data for a range of shellfish species at coastal sites in Western Scotland, SW England, and Northern France. We choose to focus on this data set to leverage the direct connection it offers to regulatory action levels for *Dinophysis* toxins (i.e. 160 μg/kg Okadaic Acid equivalents [OA eq.] in shellfish flesh), providing a pathway to better understanding impacts on cultivation of different shellfish species at a range of sites, in each of our study regions. To quantify long-term HAB risks, we present a statistical framework built on flexible smoothing techniques capable of evaluating spatio-temporal trends in average *Dinophysis* toxin concentrations in shellfish. This framework offers the potential to inform marine spatial planning and species selection for new shellfish mariculture businesses. We assess this potential using a cross-validation experiment to systematically predict toxin concentrations at locations from which data are withheld. We find more promising performance in Northern France and Western Scotland, compared to SW England. Despite regional differences in potential for spatial interpolation of HAB risk, temporal trend analysis for all three regions highlighted a statistically significant increase in *Dinophysis* toxin concentrations over the study period, corresponding with a period of climate warming. We discuss these results and consider how data-driven models such as ours can aid long-term forecasting and mitigation of HAB risk for shellfish mariculture. We present recommendations for optimising monitoring and modelling approaches to further improve the accuracy of long-term geospatial predictions of HAB risk.

2. Methods

2.1. Data

Standardised HAB monitoring data used in this analysis span 12 years (2008–2020) and include a total of 294 sites and at least 19 shellfish species farmed or collected in a) Western Scotland, b) SW England and c) Northern France. Specifically, our model for Western Scotland (a) included data from West and North-West Scotland, with sample collection dates ranging from 30th March 2009 to 8th July 2020; our model for SW England (b) included data from the south coast of Cornwall, Devon, and Dorset, with dates ranging from 4th July 2011 to 9th September 2020; and our model for Northern France (c) included data from the northern and north western coasts of France, with dates ranging from 18th November 2008 to 31st December 2019.

At routine monitoring points for each of the sites in the data set, HAB toxins are measured in shellfish meat each week during blooms, in conjunction with HAB species abundance measured according to Official Control methods stipulated in EU Hygiene Regulation (EC) No 853/2004 (EC, 2004).

2.2. Spatiotemporal models for assessing long-term HAB risk

We first describe our approach to inferring trends in HAB risk across broad spatial and temporal scales.

To assess long-term average HAB toxin levels, we developed statistical models within the Generalized Additive Model (GAM) framework. The idea is to define a model that can capture the spatiotemporal variation in phycotoxin levels using mathematically defined structures. For instance, seasonal cycles in HAB occurrence may vary substantially over a large geographical region, or the risk from HABs may increase or decrease at different locations due to changes in ocean regimes or climate. Capturing these spatiotemporal structures can better inform decisions about geographical expansion/diversification of mariculture. In contrast to applying a different model for each site, identifying any space-time apparent structures will provide more robust estimates of risk in each site, since data will be pooled across the sites. It will also enable the prediction of risk at times and in locations where no observations exist.

We begin by assuming that toxin samples y can be modelled using a

log-Normal distribution (i.e. $\log(y) \sim \text{Normal}(\mu, \sigma^2)$). We can then characterise systematic variability in average toxin levels by introducing spatiotemporal structures in the mean μ. These structures are defined by smooth functions of one or more predictor variables (in this case space and time). The smooth functions themselves are linear combinations of basis functions, namely regression splines (Thompson et al., 2017). For example, a smooth function $f(t)$ of time t is defined as

$$f(t) = \sum_{k=1}^{K} \theta_k b_k(t)$$

where θ_k are unknown coefficients and b_k are the basis functions (such as cubic splines or thin plate splines) (Thompson et al., 2017). When K (conventionally termed the "number of knots") is too large, function f will be too "wiggly" and thus tend to overfit (over explain) the data, which is why when such functions are estimated from the data, a smoothness constraint is imposed to prevent overfitting. This constraint is estimated objectively from the data by leveraging in-sample and out-of-sample predictive skill.

For toxin measurement $y_{t,s,i}$ observed on day $t = 1, \ldots, \text{N}$, at day-of-the-year $d(t) = 1, \ldots, 365$), at geographical location $s \in S$ and in species i (e.g. mussels) the our model is defined by:

$$\log\left(y_{t,s,i}\right) \sim \text{Normal}\left(\mu_{t,s,i}, \sigma^2\right);$$

$$\mu_{t,s,i,j} = \alpha + \beta_i + f_1(t) + f_2(d(t)) + f_3(\text{s}) + f_4(\text{s},\ d(t)) + f_5(\text{s},\text{t}),$$

where α is the intercept and β_i is the species effect that captures overall differences in the uptake of HAB toxins by different shellfish species with blue mussels acting as the baseline (intercept) species for all regions. Note that all functions $f_1, \ldots, f_5$ are centred at zero so that α is the overall mean (across all space and time) toxin for blue mussels and $\alpha + \beta_i$ is the overall mean for species i. In Section 3.3 we assess differences across species using estimates of α and β_j (and their uncertainty).

Function $f_1(\text{t})$ captures the overall temporal trend (across all space) in toxin level. It is defined using thin-plate splines (Thompson et al., 2017) with 50 knots (using a specific basis representation that imitates a 1D Gaussian process with a Matérn covariance function) (Kammann and Wand, 2003; Wood, 2017). This function captures any temporal trend that is common across all sites (e.g. effects of large scale climate and ocean regimes), and therefore contributes to pooling of the data. For instance, sites with less information can "borrow" information from more data-rich sites through this common term. Function $f_2(d(t))$ describes the overall seasonal trend across the region. This is defined as a cyclic cubic spline of calendar day (cyclic means the value is the same for day 0 and day 365) and is given 10 knots. As with $f_1(\text{t})$, pooling is achieved since this is common to all sites.

The function $f_3(\text{s})$ is two-dimensional and it defines a surface over the coordinates of the region. It captures the overall spatial structure across the whole time period and it therefore pools the data spatially (e.g. by borrowing information from nearby data-rich sites). $f_3(\text{s})$ is specified as a 2D thin-plate spline of longitude and latitude. The specific basis representation imitates a 2D Gaussian process in space with an isotropic Exponential covariance function (Wood, 2017). For this function, we specified 50 knots for SW England, 100 for Western Scotland and 100 for Northern France, reflecting the smaller geographical extent of SW England compared to Western Scotland and Northern France.

Function $f_4(\text{s},\ d(t))$ is three-dimensional and it describes a different spatial surface for each day of the year $d(t)$. This allows the seasonal cycle to be different for each site (and more generally location s). Similarly, $f_5(\text{s},\text{t})$ defines a different spatial surface for each day t, therefore allowing each site to have a different temporal trend if necessary. Both f_4 and f_5 are constructed using tensor products, which can interact different types of smooth functions together. First, f_4 interacts a 2D thin plate spline (50 knots for Northern France and Western Scotland, 25 for SW England) of latitude and longitude with a 1D cubic cyclic spline of calendar day (10 knots). The result is that spatial surfaces are similar for adjacent days and days in the year. Meanwhile, f_5 combines a 2D thin-plate spline of longitude and latitude (10 knots) with a 1D thin-plate spline (10 knots) of time. The structured nature of these interactions means that the seasonal and temporal trends are constrained to be more similar for nearby sites, while the spatial surface is more similar for successional days. One can also think of f_4 and f_5 as capturing the deviations of each site from the overall trends f_2 and f_1 respectively. In choosing to specify highly structured interactions we aim to allow robust prediction at locations where there are no observations (we later assess this capability for Northern France in Section 2.5). The models are fitted to the data using the R package mgcv (Wood, 2011).

2.3. Estimating yearly changes in toxin concentrations

The spatiotemporal models described in Section 2.2 estimate the temporal signal in mean (log) toxin concentrations through 1D thin-plate splines $f_1(\text{t})$, but since these functions are non-linear it can be difficult to assess any longer-term trends. We therefore opted to fit supplementary models, for each region, which replace all terms involving time with a single simple linear effect of time (in days). The coefficients of these effects were all positive and significantly different from zero at the 0.1% level. We calculated average "year-on-year" increases by multiplying the coefficients by 365 and exponentiating the result (to invert the log transformation applied to the toxin concentrations). These increases are reported in Section 3.2.

2.4. Handling of missing values

In both the UK phycotoxin datasets, particularly for SW England, a high proportion of observations were recorded as less than the Reporting Limit '<RL' – around 80%. Reporting Limit (RL) values are higher than detection limit (DL) values; all values reported by the UK FSA represent upper confidence intervals around the final calculated toxin concentrations, and therefore provide greater protection of human health. Nonetheless, these <RL observations can provide valuable information on the periods of the time and locations where the toxin level was low. To accommodate such data in our spatiotemporal models for SW England and Western Scotland, we replaced all <RL entries with a fixed value below the lowest measured value, 10 μg/kg OA equivalents. For the French phycotoxin data, only numeric toxin values are reported (i.e. no <RL or equivalent). We therefore assume that the entire distribution of toxin concentrations had been observed (i.e. we assume there are no non-detects), noting that just under half of measurements are low values (i.e. below 10 μg/kg OA equivalents). If this assumption is wrong, for instance if non-detects have in fact already been replaced in the data with fixed low values, then our results may suffer from some bias at low toxin levels (e.g. less than 10 μg/kg OA equivalents). This bias also results from replacing <RL values with 10 μg/kg OA equivalents in the UK datasets. However, we do not believe bias at low toxin levels is a significant issue for our aim to quantify risks from HABs, in the context of which precisely capturing low toxin levels is of low importance compared to reliably differentiating between low and high toxin levels.

Some toxin measurements (approximately 0.78%) from shellfish samples in the data set for Northern France had recorded longitude and latitude values placing them over land, so we excluded them from the model.

2.5. Cross-validation experiment

To assess the capability of our approach for predicting HAB risk at locations with no data, we carried out a k-fold cross-validation experiment. The parameter k represents the number of groups the data are split into. We choose $k = 10$, as this value has been shown empirically to be reliable in terms of the bias and variance of out-of-sample errors (James

et al., 2013). For each region separately, the procedure was as follows:

1 Shuffle the harvesting sites into a random order.
2 Split the shuffled sites into 10 groups of approximately equal size.
3 For each group.

a exclude all the data for the sites in that group;
b fit the spatiotemporal model to data from the remaining 9 groups;
c use the fitted model to predict the log-toxin values for the excluded data points.

Here, we believe removing data for whole sites at a time is a more stringent implementation of k-fold cross-validation than just splitting the data into k groups irrespective of site, but better suited to assessing the capability of the model for assessing risk at locations without data.

We then computed the mean predicted log toxin value for each harvesting site (across the whole time period of the study), as an indicator of *Dinophysis* HAB risk. To assess out-of-sample prediction performance across regions, we focus on relative HAB risk, since for Western Scotland and SW England a high percentage of available toxin measurements were below the reporting limit (Section 2.4), making assessment of absolute toxin level prediction less meaningful. We quantify relative risk by comparing the *ranked* (ascending) observed and predicted (out-of-sample) mean toxin levels at each site. We summarise performance in each region using scatter plots of observed versus predicted ranks and computing the correlation of the observed and predicted ranks (Spearman's ρ). These results are presented in Section 3.5.

For Northern France, where we did not need to replace any values below the reporting limit (Section 2.4), we also assess prediction of absolute HAB risk at unseen locations. In this case, we summarise performance by computing the Pearson correlation value of the predicted versus observed mean log toxin values at each site (excluding sites with less than 20 toxin measurements). These results are also reported in Section 3.5.

3. Results

3.1. Overall spatial distribution of Dinophysis toxin concentrations

Estimates of $f_3(s)$ (see Section 2.2), which can be interpreted as the long-term average HAB toxicity, can be used to indicate persistent or frequently recurring blooms and can therefore inform decisions about suitable locations for shellfish farming away from these hotspots. Here we plot the estimates at the log-scale so that e.g. a location with a plotted value of 2 has an $\exp(2) \approx 7.4$ times higher mean toxin concentration than a location with a plotted value of 0. Fig. 1 shows the estimates for South-West England, while Fig. 2a shows Western Scotland and Fig. 2b shows Northern France.

In SW England, *Dinophysis* HAB hotspots (highlighted in yellow) include shellfish production areas mainly for blue mussels (*Mytilus edulis*) and Pacific oysters (*Magallana gigas*) in Falmouth Bay and St Austell Bay to the west (Fig. 1). In Scotland, regularly recurring HAB hotspots are located in the relatively sheltered Clyde Sea and along the more indented NW coast (Fig. 2a), impacting mainly on the farming of blue mussels and Pacific oysters.

Shellfish production sites in Northern France are depicted for the north and west coasts (Fig. 2b); HAB hotspots comprise wild beds, including offshore king scallop (*Pecten maximus*) beds in the English Channel, as well as embayments along the north and west coast, which contain wild and cultivated King scallops (*Pecten maximus*) and Queen scallops (*Aequipecten opercularis*), Pacific oysters and blue mussels.

3.2. Seasonal variation in Dinophysis toxin concentrations

A clear understanding of systematic seasonal variability, or in other words what we expect the average toxin levels in shellfish to be at different times of the year, is essential for developing a sustainable mariculture industry, i.e. for site selection and subsequent mitigation of HAB impacts through optimal harvest scheduling and supply chain management. Our models for Northern France, Western Scotland and SW England estimate the overall temporal ($f_1(t)$) and seasonal signals ($f_2(d(t))$) in *Dinophysis* toxin concentrations at a regional level for each country, as illustrated in Fig. 3.

In all 3 regions, toxin concentrations indicate an increasing trend over the period 2012–2020, as shown by linear trend lines (solid lines). In SW England, each peak in the temporal signal was higher than the last, interrupted temporarily in 2017 when concentrations fell substantially (Fig. 3a). On average, mean toxin concentrations increased by 4.2% year-on-year in Northern France, by 2.7% year-on-year in SW England, and by 2.9% year-on-year in Western Scotland, within the time period and areas covered by the data (Section 2.1, also see Section 2.3 for details on computing year-on-year changes). Maximum toxin concentrations are recorded from all three data series in 2018. The seasonally averaged signals for Scotland and SW England are similar, displaying an approximately symmetrical shape with single peaks occurring around July-August, while the highest toxin concentrations for Northern France occur around June, followed by a further small increase/inflection in the Autumn (Fig. 3b).

Fig. 4 examines space-season interactions ($f_4(s, d(t))$) for Northern France, where the model shows that peak toxicity occurs in July along France's Atlantic coast, while on the north coast of France the peak occurs later in the summer and early autumn.

3.3. Assessing vulnerability of shellfish species for spatial planning

The model described in Section 2.2 also estimates the average

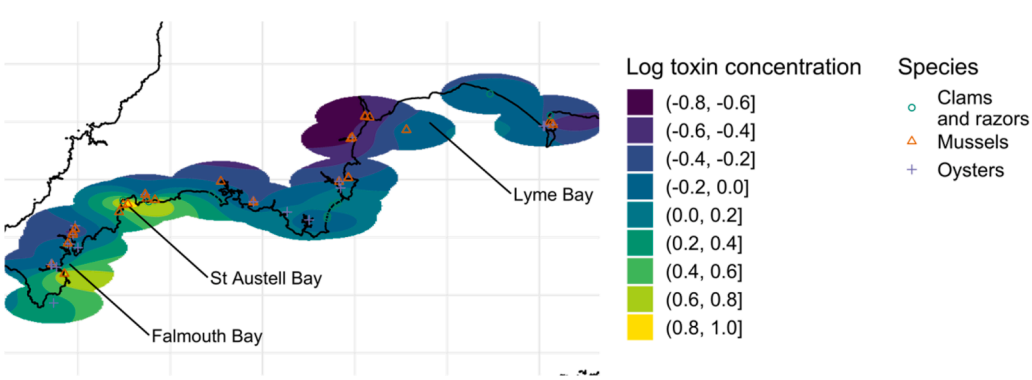

Fig. 1. Estimated spatial variation in log Dinophysis toxin concentrations in shellfish in South-West England. Symbols indicate location of shellfish sites and species monitored for Dinophysis toxins.

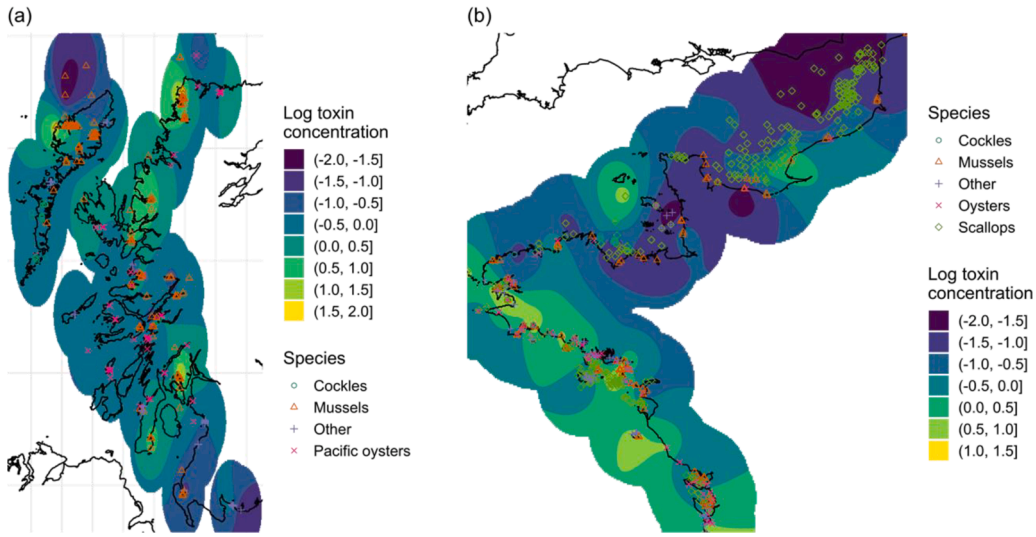

Fig. 2. Estimated spatial variation in log *Dinophysis* toxin concentrations in shellfish in (a) the west coast of Scotland and (b) the north and west coasts of France. Symbols indicate location of shellfish sites and species monitored for *Dinophysis* toxins.

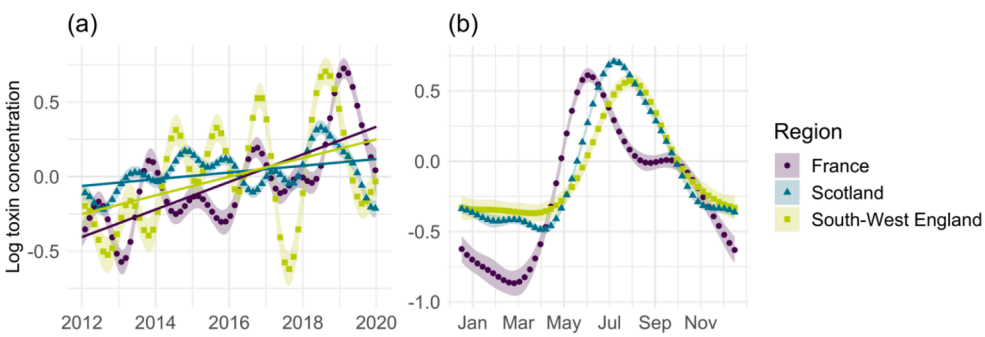

Fig. 3. Estimated (a) temporal and (b) seasonal signals (plotted as shapes) in the mean (log) *Dinophysis* toxin concentration in shellfish in each study region (from the spatial/main effects model), with 95% confidence intervals (shaded areas). Solid coloured lines are linear trend lines for each region. Note that (a) and (b) respectively show the separated temporal and seasonal signals. One would have to combine these two effects to see the annual maximum and minimum toxin concentrations occurring at the expected time of year.

Dinophysis toxin concentration in different shellfish species relative to the average toxin concentration in blue mussels (*Mytilus edulis*). By considering both the point estimates and uncertainty (95% confidence intervals) shown in Fig. 5, we can draw tentative conclusions about the risks associated with the farming of different shellfish species. We only compared estimates for species with at least 200 data points in total across all regions. Several species exhibited similar average levels of intoxication to mussels, including Dog cockles (*Glycymeris glycymeris*), King scallops (*Pecten maximus*) and Queen scallops (*Aequipecten opercularis*). Meanwhile, Manila clams (*Venerupis philippinarum*), Pacific oysters (*Magallana gigas*) and common cockles (*Cerastoderma edule*) in Northern France exhibited substantially lower average toxin concentrations than blue mussels. Both native Atlantic oysters (*Ostrea edulis*) in Scotland and abrupt wedge shell clams (*Donax trunculus*) in Northern France exhibited substantially higher average toxin levels compared to blue mussels.

Estimating the relative *Dinophysis* toxin concentrations for different shellfish species provides the opportunity to assess the potential counterfactual change in HAB intoxication risk associated with the farming of alternative species at existing sites. For example, for Antifer ponton pêche on France's north coast, *Dinophysis* toxin concentrations in blue mussels, consistently exceeded the action level (160 μg/kg OA eq., requiring harvesting closure) between 2010 and 2018, often for a month at a time (Fig. 6). Our model suggests that, if it were feasible for Pacific oysters or common cockles to be farmed here instead, the toxin concentration, and consequently the risk of disruption to harvesting, could be substantially reduced (Fig. 6). Meanwhile, harvesting of abrupt wedge shell clams here could suffer from longer periods of enforced closure on average.

3.4. Assessing the importance of factors contributing to HAB risk

To assess the relative importance of the different functions that define the mean log-concentration μ (section 2.2), we can compute the percentage of the variance of the toxin measurements y explained by each term. For Northern France, the most important term was the spatial function f_3 (21%), followed by the species effect (16%) – reflecting the wide variety of species being cultivated in France – and the space-season interaction f_4 (12%). In SW England, the spatial function f_3 was also most important (16%), followed by the space-season interaction f_4 (13%) and then the function of time f_1 (12%) – the latter reflecting the significant interannual variability for SW England as seen in Fig. 3. Finally, in Scotland the spatial and space-season terms were tied as the most important (19% each), followed by the seasonal function f_2(17%).

3.5. Predicting HAB risk at unseen locations

Models developed within our proposed spatiotemporal framework can be used to predict *Dinophysis* toxin levels at times and locations where no data is available, since the fitted smooth functions are continuous in space, time, and time of year. In general, we can predict average log-toxin concentrations for any date within the data period and for any new or existing location in terms of longitude and latitude. From these predictions, we can derive measures of HAB risk, such as monthly means, which can be interpreted in absolute terms or in relative terms, e.g. against average levels at existing harvesting locations.

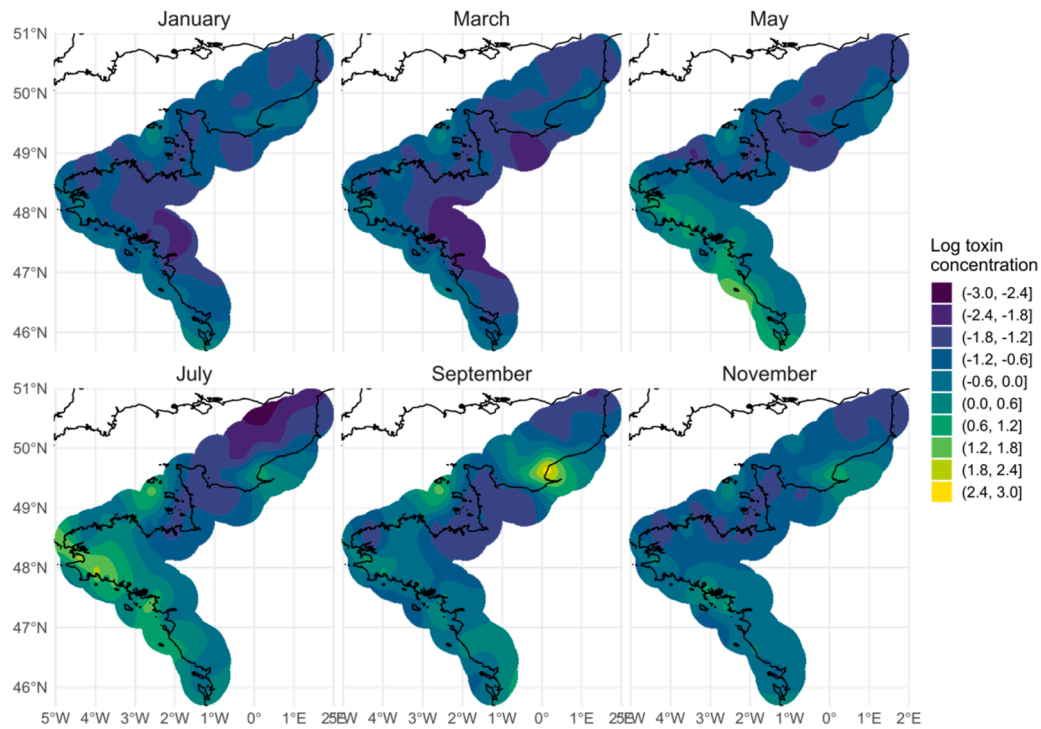

Fig. 4. Estimated average (relative) *Dinophysis* toxin levels in shellfish in Northern France at different times of the year, from the spatiotemporal model for French *Dinophysis* toxin data.

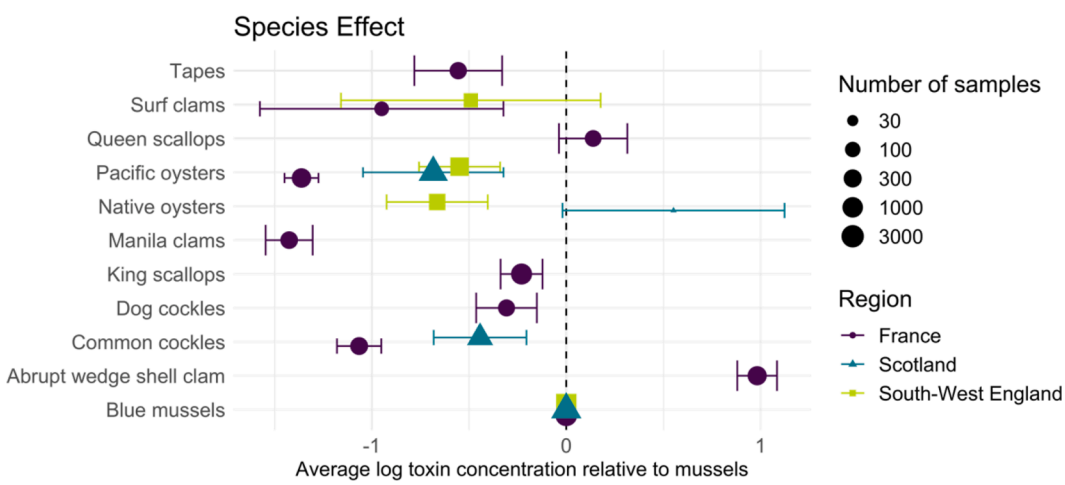

Fig. 5. Estimated species effect (variation in Dinophysis toxin concentrations across shellfish species) from the long-term spatial risk models for Northern France, Western Scotland, and South-West England. Effects are shown relative to the point estimates for blue mussels (Mytilus edulis) for each region. Points show point estimates and error bars show 95% uncertainty intervals. The size of the point relates to the number of unique observations for each species, separately for each region.

We assessed this capability by systematically excluding data for groups of harvesting sites and making out-of-sample predictions, following a 10-fold cross-validation procedure described in Section 2.5. As explained in Section 2.5, for an even comparison of performance across regions, we focussed on predicting harvesting site 'toxicity' ranks (in ascending order, as measured by mean toxin values in shellfish at each site), as an indicator of relative *Dinoyphysis* HAB risk.

Fig. 7 shows scatter plots of observed versus out-of-sample predicted toxicity ranks for harvesting sites in each region. The closer points are to the black diagonal line, the more accurately their toxicity rank has been predicted. The predicted ranks correspond most closely with the observed ranks in Northern France (correlation 0.78, $n = 203$), with only a few sites far from the diagonal. In Western Scotland (correlation 0.64, $n = 69$), most sites follow the diagonal quite closely, with a handful of sites far from the diagonal. In SW England, which had the least data and the highest percentage of <RL values, there is a weaker correspondence between predicted and observed ranks (correlation 0.34, $n = 22$).

Finally, we also assessed prediction of absolute *Dinophysis* HAB risk at unseen locations in Northern France, by comparing predicted versus observed mean (log) toxin measurements in shellfish at each harvesting site. A strong correspondence was seen between the predicted and observed site averages (correlation 0.82), not substantially worse than for prediction of seen sites in the model with all data included (correlation 0.91, $n = 118$).

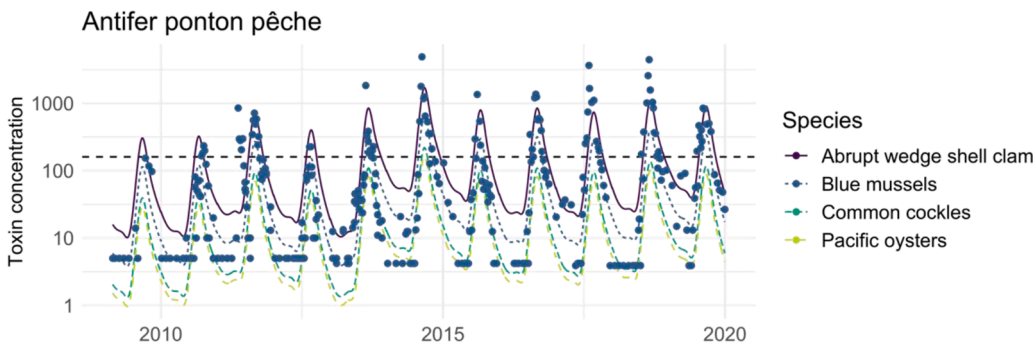

Fig. 6. Estimated mean toxin concentration at Antifer ponton pêche for different species (lines). Points show the observed concentrations from blue mussel samples and the horizontal dashed line shows the toxin threshold above which harvesting is shut down.

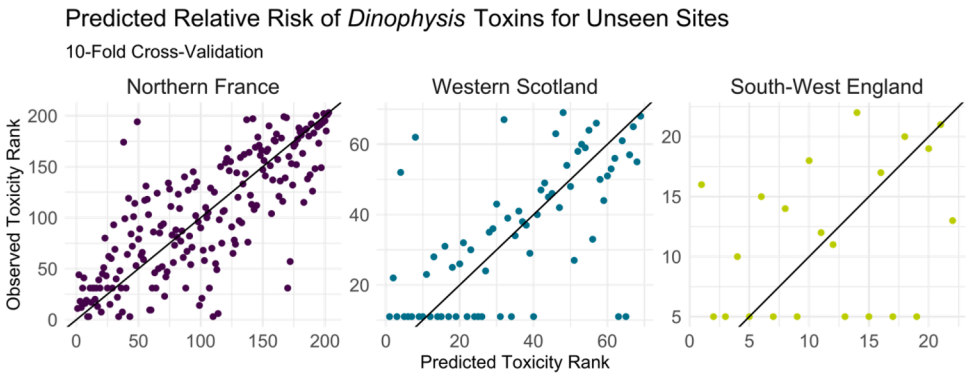

Fig. 7. Predicted versus observed rank of out-of-sample (unseen) harvesting sites by mean *Dinophysis* toxin measurement in shellfish, for each region, from the 10-fold cross-validation experiment.

4. Discussion

4.1. Spatiotemporal assessment of HAB risk for sustainable mariculture development and operation

Better quantifying and predicting the risk of HABs intoxicating shellfish is valuable for the operation and growth of shellfish mariculture. This paper describes and evaluates the use of data-driven approaches for making spatial and temporal predictions of HAB risk based on existing, standardised HAB monitoring data for shellfish production sites in Western Scotland, SW England and Northern France. The same approaches could be developed and applied elsewhere in Northern Europe, including Spain and Portugal and other key shellfish harvesting areas, which are plagued by *Dinophysis* HABs and subject to rigourous 'Official Control' monitoring.

Our spatiotemporal models predict average annual HAB toxin risk maps for a variety of shellfish species across established, new and data-poor sites. As marine aquaculture looks to expand, these spatial maps of average HAB risk could support marine spatial planning by identifying: a) recurring HAB hotspots versus lower risk areas and b) opportunities for optimisation of existing mariculture, through changes in farmed shellfish species (informed by the weight of evidence across regions) and seasonal harvesting schedules.

Western areas in SW England sheltered from prevailing SW winds by headlands were shown to have greatest risk from *Dinophysis* spp.; these areas have also been shown to undergo rapid sea surface warming and prolonged thermal stratification during the summer (Pingree et al., 1975; 1983; Brown et al. 2022). In contrast, shellfish sites in sea lochs with a southwest facing aspect along the Western coast of Scotland appear to be at greatest risk, due to the potential advection of *Dinophysis* HABs from offshore (Gianella et al., 2021).

The French dataset showed earlier onset of *Dinophysis* blooms and elevated toxin levels in shellfish on the west coast compared to the North coast (reported previously by Belin et al., 2021) and highlighted higher toxin concentrations in certain shellfish species, such as the abrupt wedge shell clam (*Donax trunculus*) compared to mussels *Mytilus* spp.). Previously mussels have been found to accumulate higher levels of diarrhetic shellfish toxins (DST), including *Dinophysis* toxins, than other species when exposed to the same algal bloom conditions (reviewed in Blanco, 2018; Swan et al., 2018). Lower levels of *Dinophysis* toxin accumulation for example in Pacific oysters or common cockles, offers the potential to strategically target the cultivation and collection of alternative, lower risk species (Fig. 6).

Much of the predictive power of our models stems from the integration of substantial (10–20 year) EU-standardised, national shellfish monitoring data. These extensive data sets are updated on a weekly (or bi-weekly) basis, allowing for continual model recalibration and assuring the validity of spatio-temporal interpolations for marine spatial planning for mariculture. Retrospective analysis of data from the most recent decade (2012–2020) demonstrated year-on-year increases in mean *Dinophysis* toxin concentrations in shellfish of 4.2% for Northern France, 2.7% for SW England, and 2.9% for Western Scotland. This period has corresponded with climate warming. For example, in 2020, annual mean sea surface temperature in near-coast waters around the UK was 11.9°C, 0.5°C above the 1981–2010 long-term average. In the most recent decade (2011–2020), sea-surface temperatures have been 0.7°C warmer than the 1961–1990 average, and nine of the ten warmest years (in terms of near-coast sea-surface temperatures) have occurred since 2002 (Kendon et al., 2020).

While our analysis does show increasing trends in all three regions, there is considerable variability between years, which makes it potentially unwise to extrapolate them beyond the latest available data. In other words, although we can predict the location of HAB hotspots based on recurrences in the last decade, longer-term multi-decadal data may

be required to more confidently project trends into the future and confirm the causal factors which drive HABs. At the same time, however, we should remain conscious that an increasingly deep historical record won't necessarily alert us to future regime changes. For example, there have been notable increases in the occurrence of HABs following major hydroclimatic regime shifts, such as the North Atlantic Oscillation switching abruptly (NAO+) and generating stronger westerly winds and milder, wetter winters since the mid-1980s in the North-East Atlantic and North Sea (Gobler et al., 2017; Raine et al., 2008; Drouard et al., 2019).

4.2. Scope for improving spatio-temporal risk assessment for HABs

Optimising spatio-temporal predictions of HABs is important for marine spatial planning for mariculture, particularly for low biomass HABs such as *Dinophysis* spp., whose progression is difficult to track by traditional microscopic methods or by high resolution Earth observation systems (Caballero et al., 2020). Our regional spatio-temporal models provide a good indication of spatial and temporal trends for *Dinophysis* toxin concentrations across routinely monitored shellfish production sites in Western Scotland, SW England and Northern France. Additional monitoring points would be beneficial and could be focused on likely HAB hotspots (e.g. southwest facing coastal areas of NW Scotland, which are more likely to intercept wind-advected blooms). It is equally important to obtain monitoring data for low-risk areas with contrasting bio-geographies and ecological niches, to confirm their potential suitability for shellfish farming. Our data-driven models could be further refined by accounting explicitly for climate variation (e.g. with latitude) and hydrodynamic factors (e.g. coastal circulation, mixing and upwelling), which differentiate regions and shellfish production sites within them.

Environmental factors could also be incorporated explicitly in the selection of shellfish fish for cultivation in a particular area, in addition to modelling the comparative uptake of HAB toxins across different shellfish within predefined regions (as illustrated here). After all, local hydrographical, water quality and climate-related conditions will ultimately govern the suitability of sites for shellfish cultivation.

Further development and refinement of our spatio-temporal model could include more detailed resolution of coastal boundary effects. Spatial structure is currently defined in our model by assuming space is a continuum, whereas for example sites either side of a peninsula may actually be more dissimilar than suggested by their Euclidian distance. Taking into account the land boundary and other geographical features could result in more reliable predictions when spatially interpolating and extrapolating, noting that the current models can produce strange estimates when predicting over land or far into the sea. Accounting for realistic structures could be achieved by using more sophisticated splines, such as soap film smoothing functions (Thompson et al., 2017), or by considering water depth as a covariate.

4.3. Scope for improving data acquisition for HAB risk assessment

Improving the spatial and temporal coverage of monitoring data for HABs and the detection of HAB toxins in shellfish (including detection levels) will clearly benefit data-driven modelling of HAB risk. This was demonstrated by data for SW England, which had fewest data points and the highest percentage of <RL values for *Dinophysis* toxin concentrations and showed far weaker correlation between out of sample predicted and observed ranks of toxin concentrations for harvesting sites in the region (Spearman's ρ=0.34, n=22), compared to Western Scotland (ρ=0.64, n=69) and Northern France (ρ=0.78, n=203). Moving towards more proactive and wider monitoring data collection to feed data-driven HAB models such as ours, currently relies on labour intensive manual sampling (e.g. weekly sampling) and highly skilled microscopic analysis of HAB cell counts, and chemical analysis of HAB toxin concentrations in shellfish which may not be feasible in every country. Nevertheless, there are several avenues for improving data collection to benefit HAB risk assessment. First, resolution of *Dinophysis* bloom dynamics could be improved by acquiring real-time data from in situ sensors capable of quantifying changing HAB toxin concentrations and HAB species abundance via automated algal cell imaging, cytometric or molecular-based methods (McPartlin et al., 2017; Bickman et al., 2018; Scholin et al., 2018; Guo et al., 2021). So far, only a limited number of real-time monitoring systems have been implemented in national or regional HAB surveillance programmes: NOAA's HAB operational forecast system for the Gulf of Mexico (Campbell et al., 2013; NOAA, 2017); the Autonomous Ocean Sampling Network including Monterey Bay (Scholin et al., 2018), the Gulf of Maine (Anderson et al., 2019), and the Hong Kong coast (Yamahara et al., 2019). There is huge scope to expand and advance HAB monitoring systems, but this work should be undertaken in tandem with the advancement of data-driven models, which can help to optimise data collection in time and space, to fill critical data gaps (e.g. for mariculture development sites that are potential HAB hotspots) and to exploit these data for forecasting and helping to mitigate the risk of HAB impacts on mariculture.

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.

Data availability

All data and model codes are available here: https://doi.org/10.5281/zenodo.7119036

Acknowledgements

The authors gratefully acknowledge funding from the European Maritime and Fisheries Fund (ENG3103); Turing Pilot Research Grant (260320) and IIB Open Innovation Project Fund (115717). T. Economou was funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 856612 and the Cyprus Government.

Author Contributions

Stoner undertook conceptual development and implementation of the statistical models, and preparation of the manuscript. The following authors also contributed to the manuscript: Economou provided expert input on the study and statistical framework design; Ashton, Torres and Brown provided expert input on climatic, oceanographic covariates. Brown conceived and coordinated the study and reporting.

Data and supplementary information

The following are available here: https://doi.org/10.5281/zenodo.7119036. All code and data used in this analysis are provided for download in a single ZIP file Code and Data.zip.

References

Anderson, D.M., et al., 2011. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. Ann. Rev. Mar. Sci. 4, 143–176. https://doi.org/10.1146/annurev-marine-120308-081121.

Anderson, C.R., et al., 2019. Scaling Up from regional case studies to a global harmful algal bloom observing system. Front. Mar. Sci. 6 https://doi.org/10.3389/fmars.2019.00250.

Belin, C., Soudant, D., Amzil, Z, 2021. Three decades of data on phytoplankton and phycotoxins on the French coast: lessons from REPHY and REPHYTOX. Harmful Algae 102, 101733. https://doi.org/10.1016/j.hal.2019.101733.

Berdalet, E., et al., 2016. Marine harmful algal blooms, human health and wellbeing: challenges and opportunities in the 21st century. J. Mar. Biol. Assoc. U. K. 96, 61–91. https://doi.org/10.1017/S0025315415001733.

Bickman, S.R., et al., 2018. An innovative portable biosensor system for the rapid detection of freshwater cyanobacterial algal bloom toxins. Environ. Sci. Technol. 52, 11691–11698. https://doi.org/10.1021/acs.est.8b02769.

Blanco, J., 2018. Accumulation of dinophysis toxins in bivalve Molluscs. Toxins 10. https://doi.org/10.3390/toxins10110453.

Brown, A.R., et al., 2020. Assessing risks and mitigating impacts of harmful algal blooms on mariculture and marine fisheries. Rev. Aquac. 12, 1663–1688. https://doi.org/10.1111/raq.12403.

Brown, A.R., et al., 2022. Harmful Algal Blooms and their impacts on shellfish mariculture follow regionally distinct patterns of water circulation in the western English Channel during the 2018 heatwave. Harmful Algae 111, 102166.

Caballero, I., et al., 2020. New capabilities of Sentinel-2A/B satellites combined with in situ data for monitoring small harmful algal blooms in complex coastal waters. Sci. Rep. 10, 8743. https://doi.org/10.1038/s41598-020-65600-1.

Campbell, L., et al., 2013. Continuous automated imaging-in-flow cytometry for detection and early warning of Karenia brevis blooms in the Gulf of Mexico. Environ. Sci. Pollut. Res. 20, 6896–6902. https://doi.org/10.1007/s11356-012-1437-4.

Costello, C., et al., 2020. The future of food from the sea. Nature 588, 95–100. https://doi.org/10.1038/s41586-020-2616-y.

Cruz, R.C., et al., 2021. A Review of Recent Machine Learning Advances for Forecasting Harmful Algal Blooms and Shellfish Contamination. J. Mar. Sci. Eng. 9 https://doi.org/10.3390/jmse9030283.

Daniels, C. et al. (2020). Supporting mariculture development: evidence for informed regulation. Policy Brief. URL (accessed June 2022) https://www.exeter.ac.uk/media/universityofexeter/research/microsites/sustainableaquaculturefutures/docs/POLICY_STATEMENT_Supporting_Mariculture_Development_.pdf.

Davidson, K., et al., 2016. Forecasting the risk of harmful algal blooms. Harmful Algae 53, 1–7. https://doi.org/10.1016/j.hal.2015.11.005.

Díaz, P.A., et al., 2019. Mesoscale dynamics and niche segregation of two dinophysis species in galician-portuguese coastal waters. Toxins 11, 37.

Drouard, M., et al., 2019. Disentangling dynamic contributions to summer 2018 anomalous weather over Europe. Geophys. Res. Lett. 46, 12537–12546. https://doi.org/10.1029/2019GL084601.

EC - European Council (2004a). Regulation (EC) No 852/2004 of the European Parliament and of the Council of 25 June 2004 on the hygiene of foodstuffs. Official Journal of the European Union, L 226/3. URL (accessed February 2020): https://eurlex.europa.eu/LexUriServ/LexUriServ.do?, uri=OJ:L:2004:226:0003:0021:EN:PDF.

EC - European Council (2004b). Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for the hygiene of foodstuffs. Off. J. Eur. Union, L 139/55. URL (accessed February 2020): https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?, ri=OJ:L:2004:139:0055:0205:en:PDF.

EC - European Council (2004c). Regulation (EC) No 854/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption. Off. J. Eur. Union, L 226/83. URL (accessed February 2020): https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?, uri=OJ:L:2004:226:0083:0127:EN:PDF.

Fernandes-Salvador, J.A., et al., 2021. Current status of forecasting toxic harmful algae for the North-east atlantic shellfish aquaculture industry. Front. Mar. Sci. 8 https://doi.org/10.3389/fmars.2021.666583.

Ferreira, J.G., et al., 2009. Analysis of coastal and offshore aquaculture: application of the FARM model to multiple systems and shellfish species. Aquaculture 292, 129–138. https://doi.org/10.1016/j.aquaculture.2009.03.039.

Food and Agriculture Organization, 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action, Rome, 2020.

Gillibrand, P.A., et al., 2016. Individual-based modelling of the development and transport of a Karenia mikimotoi bloom on the North-west European continental shelf. Harmful Algae 53, 118–134. https://doi.org/10.1016/j.hal.2015.11.011.

Gianella, F, Burrows, MT, Swan, SC, Turner, AD, Davidson, K, 2021. Temporal and spatial patterns of harmful algae affecting Scottish shellfish aquaculture. Front. Mar. Sci. 8, 785174 https://doi.org/10.3389/fmars.2021.785174.

Glibert, P.M., et al., 2014. Vulnerability of coastal ecosystems to changes in harmful algal bloom distribution in response to climate change: projections based on model analysis. Global Change Biol. 20, 3845–3858. https://doi.org/10.1111/gcb.12662 (2014).

Gobler, C.J., et al., 2017. Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans. Proc. Natl. Acad. Sci. 114, 4975. https://doi.org/10.1073/pnas.1619575114.

Grasso, I., et al., 2019. The hunt for red tides: deep learning algorithm forecasts shellfish toxicity at site scales in coastal Maine. Ecosphere 10, e02960. https://doi.org/10.1002/ecs2.2960.

Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP) (2001). Protecting the Oceans from Land-based Activities: land-based sources and activities affecting the quality and uses of the marine, coastal and associated freshwater environment. Reports and Studies No. 71.11., 168p, URL (accessed June 2022), http://www.gesamp.org/publications/protecting-the-oceans-from-land-based-activities.

Guo, J., Ma, Y., Lee, J.H.W., 2021. Real-time automated identification of algal bloom species for fisheries management in subtropical coastal waters. J. Hydroenviron. Res. 36, 1–32. https://doi.org/10.1016/j.jher.2021.03.002.

Hallegraeff, G.M., et al., 2021. Perceived global increase in algal blooms is attributable to intensified monitoring and emerging bloom impacts. Commun. Earth Environ. 2, 117. https://doi.org/10.1038/s43247-021-00178-8.

James, G., et al., 2013. An Introduction to Statistical Learning: with Applications in R. Publ. Springer, New York.

Kammann, E.E., Wand, M.P., 2003. Geoadditive models. J. R. Stat. Soc. Ser. C Appl. Stat. 52, 1–18. https://doi.org/10.1111/1467-9876.00385.

Karasiewicz, S., et al., 2020. Harmful algae niche responses to environmental and community variation along the French coast. Harmful Algae 93, 101785. https://doi.org/10.1016/j.hal.2020.101785.

Kendon, 2020. State of the UK Climate 2020. Int. J. Climatol. 41, 1–76. Supplement S2.

Lee, J.H.W., et al., 2003. Neural network modelling of coastal algal blooms. Ecol. Modell. 159, 179–201. https://doi.org/10.1016/S0304-3800(02)00281-8.

Manfrin, C., et al., 2012. Physiological and molecular responses of bivalves to toxic dinoflagellates. Invertebrate Surviv. J. 9, 184–199.

Mardones, J.I. et al. (2020). "Estimating and mitigating the economic costs of harmful algal blooms on commercial and recreational shellfish harvesters," in GlobalHAB. Evaluating, reducing and mitigating the cost of harmful algal blooms: a compendium of case studies. ed. V. L. Trainer (Sidney, BC: PICES Scientific Report), 59, 66–83.

McPartlin, D.A., et al., 2017. Biosensors for the monitoring of harmful algal blooms. Curr. Opin. Biotechnol. 45, 164–169. https://doi.org/10.1016/j.copbio.2017.02.018.

National Oceanographic and Atmospheric Administration NOAA (2017). Gulf of Mexico Harmful Algal Bloom Forecast, Tides and Currents, https://tidesandcurrents.noaa.gov/hab/gomx.html.

Pingree, R.D., et al., 1975. Summer phytoplankton blooms and red tides along tidal fronts in the approaches to the English Channel. Nature 258, 672–677. https://doi.org/10.1038/258672a0.

Pingree, R.D., Mardell, G.T., Maddock, L., 1983. A marginal front in Lyme Bay. J. Mar. Biol. Assoc. U. K. 63, 9–15. https://doi.org/10.1017/S0025315400049754.

Raine, R. et al. (2008). Harmful algal blooms in marine climate change impacts annual report card 2007–2008.

Reguera, B., et al., 2014. Dinophysis toxins: causative organisms, distribution and fate in shellfish. Mar. Drugs 12, 394–461.

Schmidt, W., et al., 2018. A generic approach for the development of short-term predictions of Escherichia coli and biotoxins in shellfish. Aquac. Environ. Interact. 10, 173–185. https://doi.org/10.3354/aei00265.

Scholin, C., et al., 2018. The quest to develop Ecogenomic sensors: a 25-Year history of the environmental sample processor (ESP) as a case study. Oceanography 30, 100–113. https://doi.org/10.5670/oceanog.2017.427.

Swan, S.C., et al., 2018. Dinophysis acuta in Scottish coastal waters and its influence on diarrhetic shellfish Toxin profiles. Toxins 10, 399. https://doi.org/10.3390/toxins10100399.

Taranu, Z.E., et al., 2017. Predicting microcystin concentrations in lakes and reservoirs at a continental scale: a new framework for modelling an important health risk factor. Global Ecol. Biogeogr. 26, 625–637. https://doi.org/10.1111/geb.12569.

Thompson, A., et al., 2017. Satellites to seafloor: Toward fully autonomous ocean sampling. Oceanography 30, 160–168.

Trainer, V.L., et al., 2020. GlobalHAB: Evaluating, reducing and mitigating the cost of Harmful Algal Blooms: a Compendium of case studies. PICES Press 28, 30–32.

van der Schatte Olivier, A., et al., 2020. A global review of the ecosystem services provided by bivalve aquaculture. Rev. Aquac. 12, 3–25. https://doi.org/10.1111/raq.12301.

Weisberg, R.H., et al., 2019. The coastal ocean circulation influence on the 2018 west Florida shelf k. brevis red tide bloom. J. Geophys. Res.: Oceans 124, 2501–2512. https://doi.org/10.1029/2018JC014887.

Wells, M.L., et al., 2015. Harmful algal blooms and climate change: learning from the past and present to forecast the future. Harmful Algae 49, 68–93. https://doi.org/10.1016/j.hal.2015.07.009.

Wells, M.L., et al., 2020. Future HAB science: directions and challenges in a changing climate. Harmful Algae 91, 101632. https://doi.org/10.1016/j.hal.2019.101632.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, Second Edition. Chapman and Hall/CRC Press, New York, 496 pages.

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x.

Yamahara, K.M., et al., 2019. In situ autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. Front. Mar. Sci. 6 https://doi.org/10.3389/fmars.2019.00373.