Ecological Informatics 60 (2020) 101154

Contents lists available at ScienceDirect

Ecological Informatics

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

Long-term trends in herring growth primarily linked to temperature by gradient boosting regression trees

Olga Lyashevska[a,*], Clementine Harma[a], Cóilín Minto[a], Maurice Clarke[b], Deirdre Brophy[a]

[a] Marine and Freshwater Research Centre (MFRC), Galway-Mayo Institute of Technology (GMIT), Dublin Road, Galway, Ireland
[b] Fisheries Ecosystems Advisory Services (FEAS), The Marine Institute (MI), Rinville, Oranmore, Co. Galway, Ireland

ARTICLE INFO

Keywords:
Herring
Clupea harengus
Growth
Time-series
Supervised machine learning
Gradient boosting regression trees
Multiple drivers
Interactions

ABSTRACT

Environmental change and fishing activity can produce directional trends in exploited fish populations with consequences for stock productivity. For herring in the Celtic Sea, size at age has been in steady decline since the mid 1980's. In the neighbouring herring stock off the Northwest coast of Ireland, reductions in size at age are noted after 1990. Here, gradient boosting regression trees were used to investigate trends in extended time series (1959–2012) of length-at-age across both populations and to identify important variables associated with the observed declines in size. The predominant signal detected was a non-linear negative relationship between adult size and mean Sea Surface Temperature during the first growing season. Herring length was negatively correlated with the Atlantic Multidecadal Oscillation. Weaker associations with indicators of food availability and population size were also detected. Across both populations a marked decline in length was observed at the upper end of the temperature range (~14°C in the Celtic Sea and ~13°C in the Northwest). Declines in length and associations with temperature were more pronounced in the Celtic Sea population which may be vulnerable to increasing sea temperatures due to its position at the southern limit of the species distribution.

1. Introduction

Identifying drivers of change in complex ecological systems and quantifying their effects is a difficult task. Ecological systems are typically influenced by multiple drivers that may combine cumulatively or interactively (Crain et al., 2008) and often exert threshold or non-linear responses (Griffen et al., 2016; Sugihara and May, 1990). These drivers may directly impact the physiology of individual organisms or may exert indirect ecological impacts, for example via trophic interactions (Koenigstein et al., 2016). Attempts to explain biological responses of fish populations using empirical data must therefore look beyond correlation with individual drivers account for additive and multiplicative effects while considering underlying mechanisms.

In fish populations, biological responses to external pressures are often manifested through changes in growth rate. Directional changes in growth have been observed across many exploited fish populations (Baudron et al., 2011; Neuheimer et al., 2011; Neuheimer and Taggart, 2010; van Walraven et al., 2010) and variously attributed to changes in the physical environment (typically temperature), food availability, population density and the selective effects of fishing (Audzijonyte et al., 2016; Conover and Munch, 2002; Law, 2000; Swain et al., 2007). Declines in growth lead to lower overall productivity via influences on survival, recruitment and fecundity (Brander, 2007). This creates an imperative to interrogate temporal change in growth, determine the combined effects of various potential contributing factors, and to improve understanding of likely future responses to the combined influence of climate and fishing (Perry et al., 2010).

Decreases in size at age over time have been observed in the Celtic Sea (CS) (ICES divisions divisions VIIa South, VIIg, VIIh and VIIj) and the North-West of Ireland (NW) (ICES divisions VIa South and VIIb-c) herring populations (Harma et al., 2012; Lynch, 2011). In the Celtic Sea, there has been a steady and pronounced decline in size at age since the mid 1980's. This has been highlighted as a significant concern for management due to impacts on stock yield (ISEC, 2019). Stock biomass

Acronyms: CS, Celtic Sea; NW, North-West of Ireland; ICES, International Council for the Exploration of the Sea; SST, Sea Surface Temperature; AMO, Atlantic Multi-decadal Oscillation; NAO, North Atlantic Oscillation; GBRT, gradient boosting regression trees; CPR, continuous plankton recorder survey; cfin, *Calanus finmarchicus*; chel, *Calanus helagolandicus*; Fbar, mean fishing mortality; SSB, spanning stock biomass; recr, number of recruits; cumF, mean estimated fishing mortality in the year of capture; totalN, total number; sal, salinity; Fcum, mean lifetime fishing mortality; CV, coefficient of variation; RI, randomisation by individuals; RY, randomisation by years; MSE, mean squared error; VI, variable importance; GAM, generalised additive model; GLM, generalised linear model

* Corresponding author.
E-mail address: olga.lyashevska@gmit.ie (O. Lyashevska).

https://doi.org/10.1016/j.ecoinf.2020.101154
Received 26 April 2020; Received in revised form 15 July 2020; Accepted 13 August 2020
Available online 01 September 2020

fell to an all-time low in 2019 and is currently below maximum sustainable yield targets. In the Northwest population long-term trends are less pronounced than in the Celtic Sea, however, size at age has decreased in recent years relative to the 1990's. Biomass and recruitment are at very low levels and the stock is considered to be in a state of collapse (ISEC, 2019). Improved knowledge of the dynamics of these stocks and the influence of environmental change is therefore of critical importance.

The increased availability of complex environmental datasets has stimulated interest in using machine learning techniques to explain patterns in ecological data (Olden et al., 2008; Peters et al., 2014). These developments are accompanied by ongoing debate about the relative merits of hypothesis-driven versus data-intensive scenarios (Elliott et al., 2016). However, when carefully curated, machine learning approaches can complement hypotheses-based research to elucidate complex non-linear relationships in ecological systems (Kelling et al., 2009; Muttil and Chau, 2007; Peters et al., 2014). Carefully supervised boosted regression trees are gaining favour in the ecological literature due to their superior predictive ability (Cameron et al., 2014; Escobar-Flores et al., 2013; Franklin et al., 2013; Leathwick et al., 2006; Maloney et al., 2012). For example, Leathwick et al. (2006) demonstrated that boosted regression trees improved predictive performance relative to generalised additive models when applied to the analysis of fish species richness in relation to environmental variables. Furthermore, boosting approaches are more effective than other tree-based methods since they allow for combining many simple models to produce a powerful model (Hastie et al., 2009). Previous studies have used the boosted regression trees to gain insight into how fish populations and communities respond to environmental change via changes in abundance and diversity (Froeschke and Froeschke, 2016; Pittman et al., 2009; Trigal and Degerman, 2015). However, few studies have used this approach to model individual level responses such as growth.

Here, we apply boosted regression trees to investigate a marked decline in herring growth in relation to potential environmental and fishery-related explanatory variables. We present an approach whereby all variables are chosen based on knowledge of the species' ecology and observed correlations.

Our objectives are to: 1) disentangle effects of multiple drivers on herring growth; 2) demonstrate the advantages of combining a machine learning approach with ecological knowledge to identify the relative importance of various exogenous variables in a dynamic system. The advantages of this approach will be demonstrated using a unique long time-series of biological data.

2. Materials and methods

2.1. Study area and data

Detailed biological and fisheries data have been collected from commercial herring landings since 1959 as part of Ireland's national fisheries monitoring programme. The data used in this study were from catches taken in the CS (ICES divisions VIIJ, VIIg and VIIaS) and the NW (ICES division VIaS) (Fig. 1). Samples were primarily from mid-water trawl catches (96% of the total).

From spatially and temporally representative samples of the landings, 50–100 herring were taken and biological characteristics were recorded (length to the nearest half centimetre interval, weight in grams [only after 1975], sex, maturity stage of the gonads and age). Total length of herring was reported to the nearest half centimetre below. To represent the midpoint of the size-class and to ensure consistency with the analyses performed in the assessment of herring (ICES, 2016), 0.25 cm were automatically added to each individual value. Age groups were determined using counts of winter rings in otoliths according to standard ageing protocols. Fish for which no age estimate were available (e.g., otoliths broken, unreadable, lost) were excluded from further analyses (ca. 3% of the total sample).

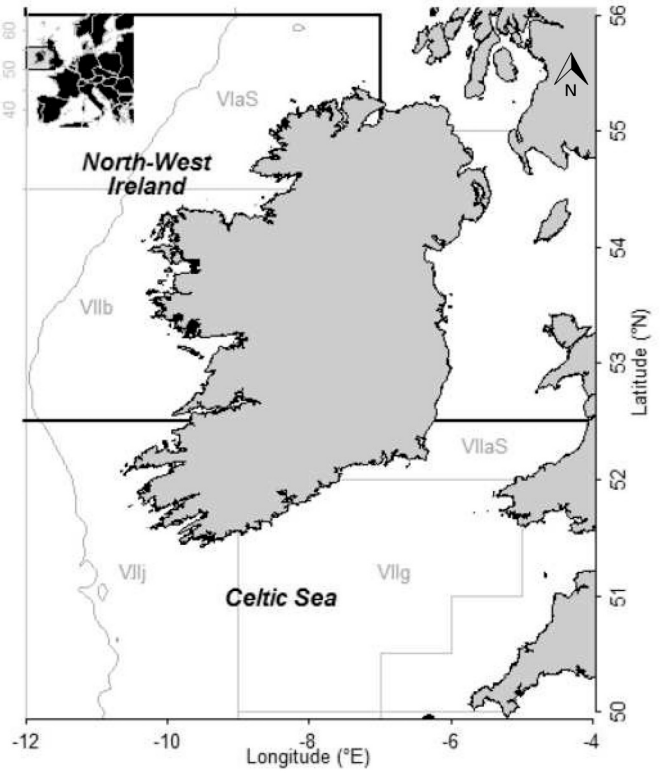


Fig. 1. Study area map showing the two Irish herring populations under study: CS for the Celtic sea populations (the ICES divisions VIIa-South, VIIg, VIIj) and NW for the North-West populations (ICES divisions VIIb and VIa-South).

Fish were assigned to year classes (yc) based on otolith ring counts as follows: $yc = ay - (r + 1)$, where ay is the assessment year and r is the winter ring count. The NW and the CS herring populations are comprised of both autumn and winter spawning components (ICES, 2015). In autumn and winter spawning herring the first translucent ring is formed in the otolith during the fish's second winter. To account for this, 1 year is added to the ring count when assigning fish to year classes.

In the CS peak spawning activity occurs between December and February and the assessment year runs from April 1st until March 31st. For this stock each year-class is referred to by the year in which spawning commences. So fish from the 2010 year class were spawned in the 4th quarter of 2010 and the first quarter of 2011. The traditional assessment year (January 1st to December 31st) applies to the NW herring stock.

The subsequent analysis was conducted using data from 3 winter ring fish because this is the youngest age group that is fully recruited to the fishery (ICES, 2016) and for which samples from the commercial fishery can be assumed to be representative of the population. The final dataset included individual records for 35,629 herring from the CS population and 22,592 herring from the NW population.

Temporal changes in length-weight relationships were examined to establish if observed changes in length of 3 ring herring coincided with changes in condition. The mean weight of a 27 cm fish was estimated from the length-weight relationship for each year (1975–2012) and area using least squares regression. The relationship between length and weight was modelled as: $ln(W) = a + b * ln(L)$, where W is fish weight in grams, L is fish length in cm and a and b are the parameters of the regression equation. The reference length of 27 cm was used because this was the average length of a 3 ring fish across the time series in both areas.

Available data for other clupeid fish populations were obtained from stock assessment reports for the purpose of comparison with

trends in the two herring populations. The closest stocks for which biological data were available were: sardine ICES Divisions VIIIc and IXa (southern Bay of Biscay and Iberian coast) and sardine in ICES Divisions VIIIabd (Bay of Biscay). For these stocks, time series of weight-at-age-3 (ICES, 2017) were plotted alongside the mean weight of 3 ring herring in CS and NW herring (ICES, 2016).

2.2. Broadscale relationships with climatic indices

As a first step in the identification of potential drivers, correlations between herring growth and two broad-scale climate indices, the Atlantic Multi-decadal Oscillation (AMO) and North Atlantic Oscillation (NAO) were investigated.

The AMO is an index of fluctuations in Sea Surface Temperature (SST) in the North Atlantic which is linearly detrended to remove the influence of anthropogenic climate change (Enfield et al., 2001). The AMO appears to alternate between positive and negative phases at a frequency of 60–80 years (Knight et al., 2006). The AMO influences marine organisms through direct effects of temperature, fluctuations in circulation patterns and changes in food availability (Edwards et al., 2013; Nye et al., 2014).

The NAO is an index of the difference in sea-level pressure difference between Iceland and either the Azores, Lisbon or Gibraltar. Cyclical fluctuations between positive and negative phases of the NAO are associated with changes in wind patterns, rainfall and temperatures (Rogers, 1997) with reported consequences for fisheries (Lehodey et al., 2006).

Proceeding from this we applied general linear models to determine if there was an association between the mean length of 3 ring herring from the CS and NW populations and the annual mean of the unsmoothed AMO index or the winter NAO index (December–March) in the first year of life. To correct for temporal autocorrelation the significance of the correlations were tested using an adjusted degrees of freedom, according to the modified Chelton method (Chelton, 1984; Pyper and Peterman, 1998), which is robust to Type I errors and less prone to Type II errors than methods based on pre-whitening or first-differencing (e.g. ARIMA) (Pyper and Peterman, 1998).

2.3. Selection of explanatory variables for GBRT analysis

While correlation analyses are useful for establishing association between a biological process and climate variability, investigation of relationships with local environmental conditions is required to build understanding of the mechanisms underlying the association (Nye et al., 2014). Here, gradient boosting regression trees (GBRT) were used to identify relationships between growth and various exogenous variables and to determine their relative strength. GBRT belong to a family of ensemble methods which employ a collection of simple additive regression model predictions that are averaged to estimate the response (Hastie et al., 2009). GBRT form a supervised machine learning algorithm, which naturally allows for complex nonlinear interactions between environmental drivers (Friedman, 2000) and do not require any assumptions on data distribution. When applying GBRT the relationship between the response and explanatory variables is not specified, instead a learning algorithm is used to find patterns in the data. Due to these flexible properties, GBRT were selected as a suitable method for identifying potential drivers of changes in growth across the two herring populations.

Separate GBRT models were constructed for each population to predict fish length based on a set of explanatory variables representing the main drivers that could contribute to variability in growth of herring (Table 1, Fig. 2). The inclusion of the two populations in the analysis facilitates a comparative investigation of their dynamics, allowing us to distinguish broad scale change from processes operating locally within the CS.

The response variable (length of 3 ring herring) reflects the combined influence of intrinsic and extrinsic factors operating over the entire life of the fish. The first growing season is a particularly important period and conditions during this critical phase can exert a lasting influence on lifetime growth trajectories (Brophy and Danilowicz, 2003; Vincenzi et al., 2008). Variation in juvenile growth is driven primarily by temperature (Ottersen and Loeng, 2000) and feeding conditions (Batten et al., 2016) with some evidence also of density dependence (Casini et al., 2006) and salinity effects (Berg et al., 2018; Rajasilta et al., 2011).

Mean monthly Reynolds Historical Reconstructed SST values (Reynolds et al., 2007) for the CS (48.5°N-52.5°N, 12.5°W-4.5°W) and NW (52.5°N-56°N, 14.5°W-7.5°W) were included as indicators of the temperature that herring were exposed to during their first growing season (average SST April–August). Among these Calanus copepods which are an important food source of juvenile herring (Huse and Toresen, 1996; Pedersen and Fossheim, 2008). Therefore, feeding conditions during the first growing season were described using indices of mean monthly abundance of *Calanus finmarchicus*,*Calanus helgolandicus* and all large copepods from three CPR standard areas (C3, C4, D4; (Richardson et al., 2006)). Salinity measurements (Ingleby and Huddleston, 2007) (average April–August) were included to reflect salinity conditions during the first growing season. To test for density dependent growth during the first year, estimates of recruitment from the stock assessment (ICES, 2014, 2015) were used as a measure of year-class strength. Total numbers of 2–9 ring fish (totalN) and spawning stock biomass (SSB) in the year of capture were also included as overall indices of population density. Two measures of fishing mortality were included to account for possible selective effects of fishing; 1) mean estimated fishing mortality (2–5 ring for CS and 3–6 ring for NW) in the year of capture Fbar and 2) mean lifetime fishing mortality for each year class of 3 ring fish Fcum. Finally, month of capture was included as an explanatory variable to account for seasonal changes in length within each year. Variability in length between sampling dates within each month was low (CV 3%) and far less than the overall percentage change in annual mean length across the time series (CS 14.6%, NW 15.8% difference between maximum and minimum length). Therefore this source of variation was not considered in the analysis.

When missing values are encountered the GBRT model will exclude all data for the corresponding time period. To avoid this, missing values were replaced using forward fill to propagate the previous value forward. The set of explanatory variables was reduced to remove highly correlated variables ($r > 0.8$; the large copepod index was excluded as it was correlated with other zooplankton indicators).

2.4. GBRT model specification

Gradient boosting regression trees (GBRT) considers additive models of the following form:

$$F_m(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \tag{1}$$

where γ_m is a learning rate and $h_m(x)$ are weak learners.

GBRT uses decision trees of fixed size as weak learners. Decision trees have a number of abilities that make them valuable for boosting, namely the ability to handle data of mixed type and the ability to model complex functions. GBRT builds the additive model in a forward stepwise fashion:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{2}$$

At each stage the weak learner $h_m(x)$ is chosen to minimize the loss function L given the current model F_{m-1} and its fit $F_{m-1}(x_i)$

$$F_m(x) = F_{m-1}(x) + \underset{h}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) - h_m(x)) \tag{3}$$

GBRT attempts to solve this minimization problem numerically via

Table 1
Description, temporal resolution, source and accessed date of datasets used as potential variables to explain herring growth variability over time.

Abbreviation	Variable	Source	Weblink
cfin C3 cfin D4 cfin C4	*Calanus finmarchicus* in areas C3, D4, C4 (mean abundance in April–August)	Sir Alister Hardy Foundation of Ocean Science CPR survey, in standard areas C3, D4, and C4. DOI https://doi.org/10.7487/2016.109.1.967	https://www.sahfos.ac.uk/
chel C3 chel D4 chel C4	*Calanus helgolandicus* in areas C3, D4, C4 (mean abundance in April–August)	Sir Alister Hardy Foundation of Ocean Science CPR survey, in standard areas C3, D4, and C4. DOI https://doi.org/10.7487/2016.109.1.967	https://www.sahfos.ac.uk/
Fbar CS	Mean fishing mortality at 2–5 rings in assessment year of capture	ICES Herring Assessment Working Group (HAWG) report 2015; Table 4.6.2.4	http://www.ices.dk
SSB CS	Spawning stock biomass in assessment year of capture		
recr CS	Number of recruits (estimated at 1 ring stage) for each year class		
Fbar NW	Mean fishing mortality at 3–6 rings in assessment year of capture	ICES HAWG report 2015; Table 5.6.12	http://www.ices.dk
SSB NW	Spawning stock biomass in assessment year of capture		
recr NW	Number of recruits (estimated at 1 ring stage) for each year class		
cumF NW	Mean lifetime fishing mortality for each year class	ICES HAWG report 2014; Table 4.6.2.13	http://www.ices.dk
totalN NW	Total numbers at age (in thousands) 2–9 ring stage, in year of capture	ICES HAWG report 2015; Table 5.6.14	
cumF CS	Mean lifetime fishing mortality for each year class	ICES HAWG report 2014; Table 4.6.2.13	http://www.ices.dk
totalN CS	Total numbers at age (in thousands) 2–9 ring stage, in year of capture	ICES HAWG report 2014; Table 4.6.2.14	
sst CS,NW	Reynolds Sea Surface Temperature (°C)	Reynolds Historical Reconstructed SST (2° x 2° resolution) as derived from the Reynolds Optimally Interpolated SST (from the Advanced Very High Resolution Radiometer, AVHRR) and in-situ observations, available from the NASA Jet Propulsion Laboratory (extracted through Hydrax/OpenDAP server), provided by Reynolds, National Climatic Data Center	https://www.esrl.noaa.gov
sal CS,NW	Salinity 5 m (PSU)	MET Office Hadley Centre EN4 quality controlled ocean data version: EN.4.2.1.	https://www.metoffice.gov.uk/hadobs/en4/download-en4-2-1.html
Month	Month, included to account for a month of capture		

Spatial resolutions for the local environmental datasets are representative of the latitude and longitude where herring populations are found in the Celtic Sea and off the North-West of Ireland areas. Abbreviations C3, D4 and C4 refer to standard areas used in the CPR.

steepest descent: The steepest descent direction is the negative gradient of the loss function evaluated at the current model F_{m-1} which can be calculated for any differentiable loss function:

$$F_m(x) = F_{m-1}(x) + \gamma_m \sum_{i=1}^{n} \nabla_F L(y_i, F_{m-1}(x_i)) \quad (4)$$

Where the step length γ_m is chosen using line search:

$$\gamma_m = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L\left(y_i, F_{m-1}(x_i) - \gamma \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)}\right) \quad (5)$$

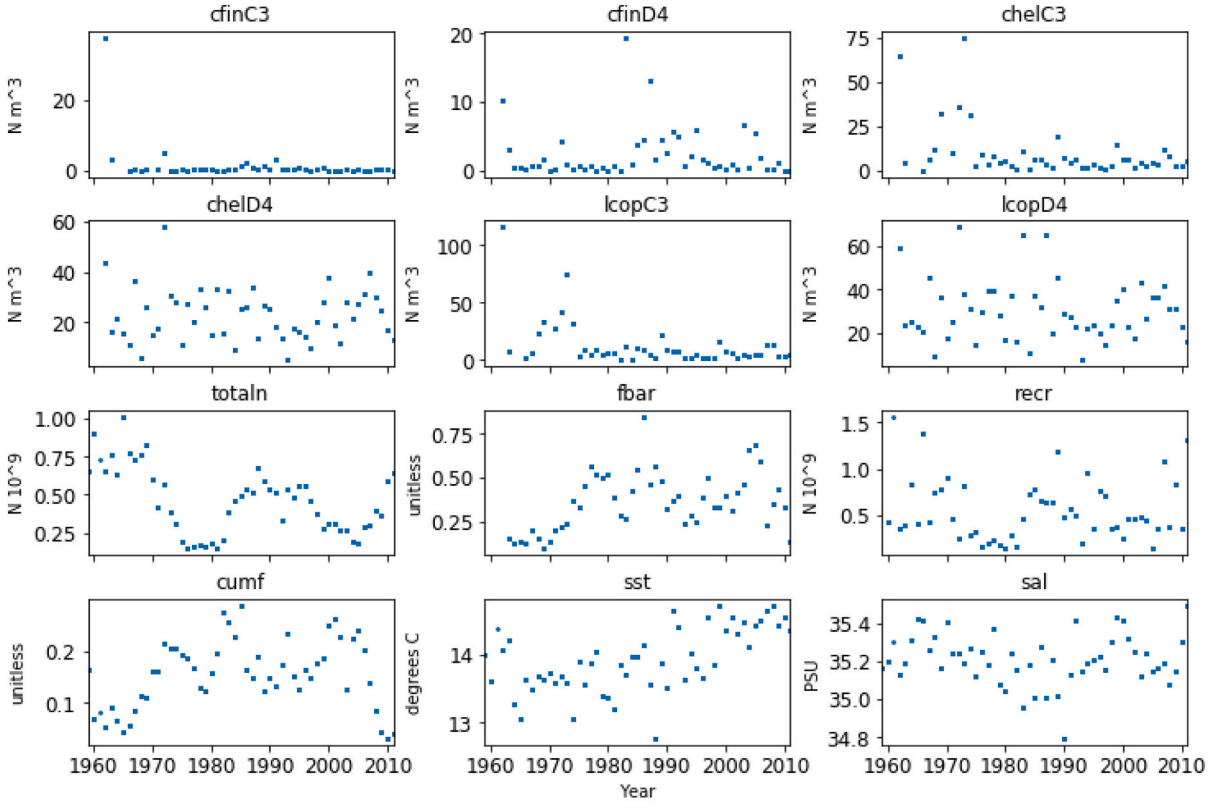
The accuracy of gradient boosting can be improved by introducing randomisation into the procedure through taking randomly selected subsets of training data at each iteration (hence stochastic gradient boosting).

For each tree the GBRT model makes splits that minimize the loss function L using the variable that contributes most in the reduction of the loss function. Variable Importance (VI) is computed as the normalized total reduction of the loss function brought by that variable and based on the number of times a variable is used in the model, weighted by the squared improvement to the model, and averaged over the entire model. The more often the variable is used for splitting, and the more it improves the model predictions, the higher the VI score.

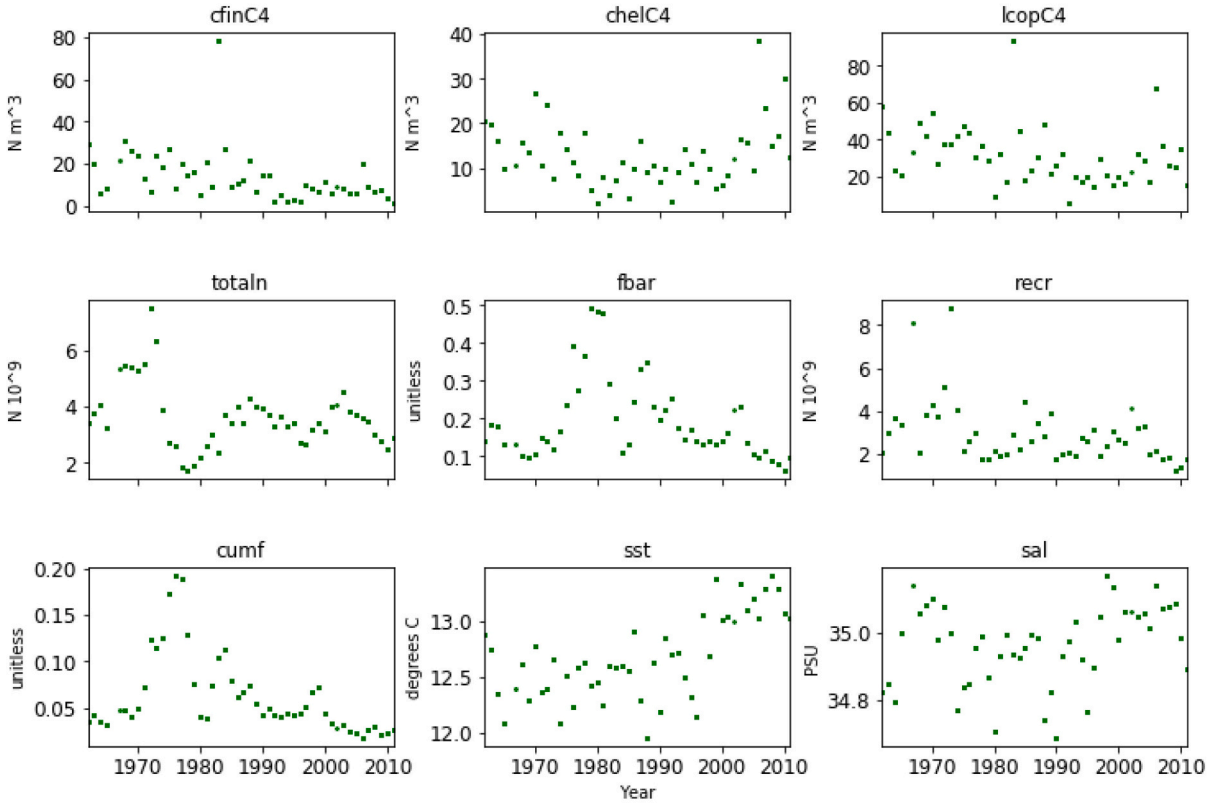
2.4.1. Data splitting by randomisation

For both populations, data were split randomly into two sets: train (50%) and test (50%). A two-part split is a general practice in machine learning, the purpose of which is to test the predictive performance of the model when presented with previously unseen data. In the present analysis, the biological datasets contained tens of thousands of individual observations each representing an individual fish. However, the environmental datasets had a much lower temporal resolution with all measurements except for month of capture aggregated annually. Therefore, if the data were split randomly at the level of individual, both sets very likely would contain individuals from all years and would not include previously unseen combinations of the explanatory variables. Such splitting could lead to overestimation of prediction capability on independent test data and may result in inadequate testing of the power of the model. To address this potential issue, two alternative randomisation approaches were employed at the data-splitting stage and the results were compared: randomisation by individuals (RI) and randomisation by years (RY). The RI data splitting was performed by randomly splitting all individual observations. This data splitting approach is generally considered optimum in machine learning as it does not introduce any systematic differences between the test and train datasets. To perform the RY data splitting, individual observations were grouped according to year of capture and years were randomly selected for inclusion in the train and test datasets. Although this may produce systematic differences between the test and train datasets, it ensures that the model is tested using previously unseen combinations of explanatory variables and is therefore a more robust test of model performance in this context.

A potential drawback of the machine learning approach is that after the model is built using the training dataset, the parameters are optimised based on how well the model performs on the test dataset. Therefore, prediction accuracy for previously unseen data may be overestimated. To perform a more robust test of prediction accuracy a



(a)



(b)

Fig. 2. Explanatory variables time series for the CS (a) and the NW (b).

three way split was also performed. The data were split (using the RI splitting approach) into train (50%), test (25%) and validation (25%) sets. The model was built, tested and optimised using the train and validation sets and the models performance was confirmed using the test set.

For the two-part split hyperparameters were tuned using grid search which performs an exhaustive search over specified parameter values for an estimator. Grid search takes a set of possible values for each hyperparameter that should be tuned and evaluates a model trained on each element of the Cartesian product of the sets. It is an exhaustive search that trains and evaluates a model for each possible combination of the hyperparameter values supplied. This algorithm automatically generates the validation sets internally. So there is no need to generate a validation set to select the best model.

For the three-part split the VI score on the validation set is used instead of the test set. Then one is allowed to tweak the values in the parameters grid to see if values that improve the score can be found. Once there is confidence that the validation score cannot be further improved via parameter tweaking (or feature engineering) one can evaluate the best model on the final test set (only once). In some instances the final test score maybe lower than the validation score. In that case the test score is taken as the most realistic evaluation of the true generalization performance of the final model.

2.4.2. Hyperparameters

Hyperparameter tuning was done manually. As a starting point the learning rate was set as low as possible and the number of iterations as high as computationally feasible (Hastie et al., 2009). The learning rate, which is also called a shrinkage parameter, determines the contribution of each tree to the model. The maximum tree depth which reflects the degree of interaction in the model is usually low. GBRT performs best using fairly shallow trees, so-called tree stumps.

For the RI models the same initial parameters were set for the analysis of the CS and NW datasets, then a grid search was performed on both datasets with a two-part split to give hyperparameters with the highest level of accuracy. These were as following: learning rate 0.05 and maximum tree depth of 4 for both the CS and the NW models. The early stopping technique was used to determine when to stop the model training to avoid overfitting. Using early stopping the number of iterations required until convergence was 120 for CS and 106 for NW. Grid search was not required for a three-part splits, and further tuning of the parameters was done manually.

For the RY models, in contrast to the RI models, the best performing models (based on MSE) were models with a slow learning rate of 0.005. Additionally, the NW model had a very shallow tree (tree depth of 2). The number of iterations was kept at 500.

The tuned hyperparameters of the final models are shown in Table 2.

2.4.3. GAM modelling of most influential parameters

The top four most influential variables from the GBRT models (as indicated by VI scores) were included in a series of generalised additive models (GAMs) with the length (of 3 ring fish) as the response variable. The purpose of this analysis was to compare the predicted relationships between the two modelling approaches. The analysis was run using the *mgcv* package, version 1.8.28 (Wood, 2011) in R, version 3.6.1 (2019-07-05). Models were fit using thin-plate regression splines. Predictors were fit using smoother terms with a maximum of 4 degrees of freedom.

Table 2
Tuned model parameters.

Model	splits	Number of iterations	Learning rate	Tree depth
RI CS	2	120	0.05	4
RI CS	3	102	0.05	4
RI NW	2	106	0.01	6
RI NW	3	144	0.05	6
RY CS	2	500	0.005	8
RY NW	2	500	0.005	2

Number of iterations was defined by early stopping. Subsample rate was 0.75% in all cases. Parameters for 2-part split were defined through grid search, whereas parameters for 3-part split were tuned manually.

3. Results

3.1. Overall trends in length, weight and condition

As previously reported (Harma et al., 2012; Lynch, 2011), mean length of 3 ring herring in the Celtic Sea showed a general increasing trend from 1960 to 1980 followed by a decline from 1981 to 2012 (Fig. 3a). Temporal changes in length of herring from the North-West of Ireland were less marked but with a slight decline in length between 1970 and 1980 (Fig. 3b).

The observed declines in size (length and weight) at age, did not coincide with a decline in body condition at a standardised size. The mean weight of a 27 cm herring over the course of the available time-series was 166.8 g in the Celtic Sea (1969–2011) (Fig. 4) and 162.4 g (1975–2011) in the North-West of Ireland. The temporal trends in the mean weight at 27 cm indicated that the observed decline in length of 3 ring herring did not coincide with a decline in condition.

Mean-weight-at age data from ICES stock assessment reports revealed recent (since the 1990's) fluctuations in weight at age across four stocks of clupeid. Trends in the mean weight-at-age 3 of sardine from commercial catches in the Southern Bay of Biscay and Iberian coast (ICES Divisions VIIIc and IXa) showed an increase in growth since 1990 which coincided with a decrease in weight of Celtic Sea and North-West of Ireland herring during the same period (Fig. 5). In contrast, mean weight-at-age of sardine in survey catches from the rest of the Bay of Biscay (ICES Divisions VIIIa, b and d) have declined since 2000.

3.2. Correlation with climate indices

The mean annual AMO index in the first year of life was negatively correlated with mean length of 3 ring herring from both the CS ($r = -0.65$, adjusted df = 10, $p < 0.001$) and NW ($r = -0.39$, adjusted df = 11, $p < 0.05$) populations. There were no significant correlations between mean length and NAO for either population ($p > 0.05$). The decline in size of Celtic Sea herring during the late 1970's and the 1980's coincided with a steady increase in the AMO index and an eventual transition from a negative to a positive phase in the mid-1990's. Earlier increases in the size of Celtic Sea herring during the 1960's and 1970's corresponded with a negative phase of the AMO and a steady decrease until the mid-1970's (Fig. 6). The increase in the AMO index in the 1970's coincided with a decline in mean length of 3 ring herring in the NW. From the mid 1990's an inverse relationship between AMO and length of NW herring was not apparent.

3.3. GBRT model performance

For the RI model the plot of MSE against number of iterations for the test data closely followed that of the train data indicating consistency in model performance across the test and train data. The difference between test MSE and train MSE was slightly greater for the NW than for the CS dataset. For the RY model there is a larger gap between test and train data. However, the difference between the train and test MSEs remained low (< 1) and prediction to independent data is unlikely to be compromised (Elith et al., 2008). Through the use of L1 regularisation technique we ensured that neither of the models was mainly affected by overfitting, which is often seen as a problem in statistical modelling. GBRT is generally superior to other methods in that regard (see e.g. comparisons with GLM, GAM and multivariate adaptive regression splines, Elith et al., 2008; Leathwick et al., 2006). Stability of

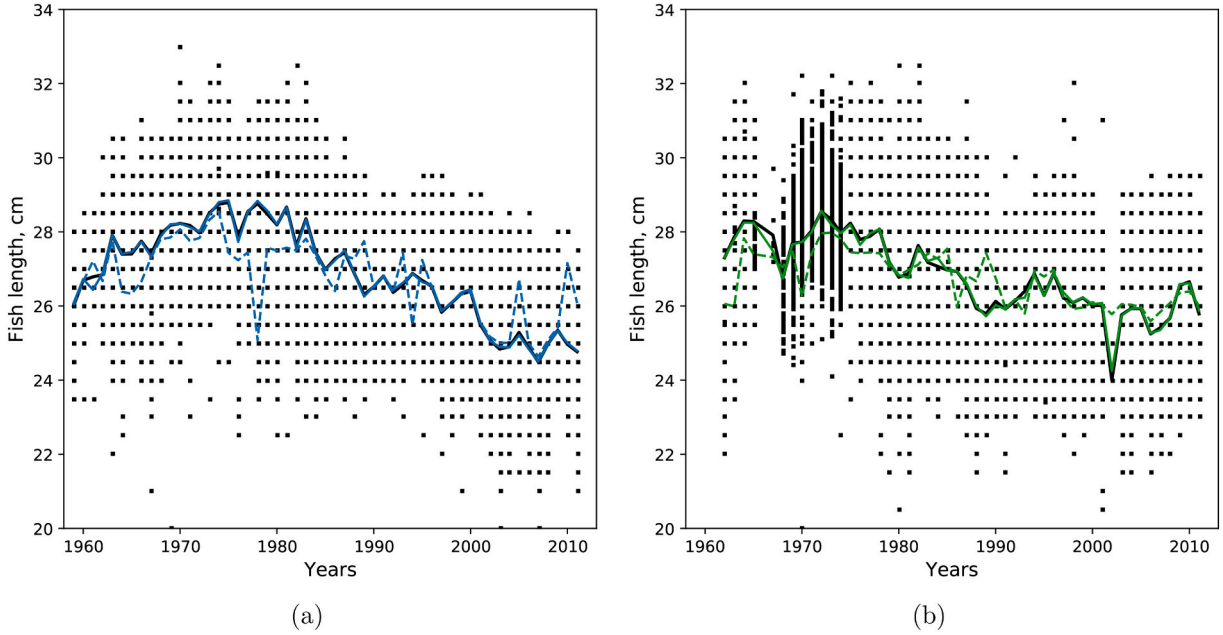


Fig. 3. Mean-length of 3 ring fish in centimetres in CS (a) and in the NW (b), cm. Measurements are rounded to the nearest 0.5 cm except for the years 1967 (a) and 1965–1974 (b). Solid line is a prediction obtained from RI model and dashed line is a prediction from RY model. Black dots are individual observations and the black line represents the observed mean length.

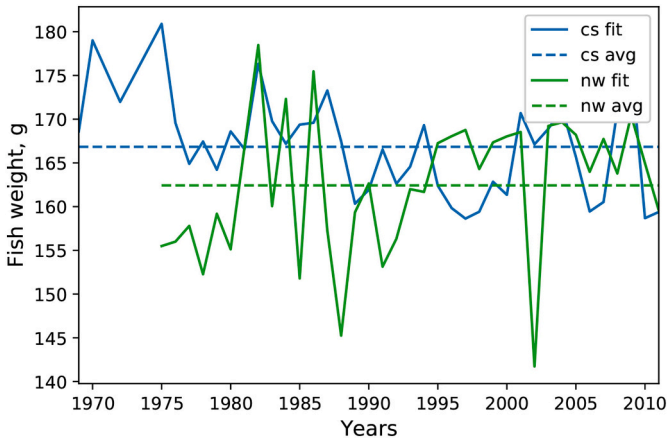


Fig. 4. Mean estimated weight of a 27 cm herring for the CS (blue solid line) and the NW (green solid line) populations. The overall means for the time series are indicated by the dashed lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

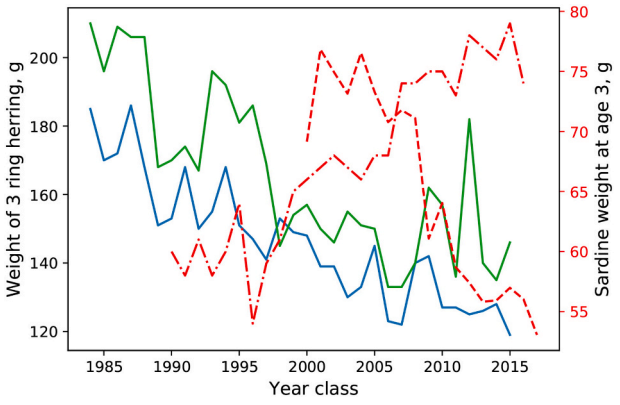


Fig. 5. Mean weight of 3 ring herring for the CS (blue solid line), the NW (green solid line), sardine in divisions VIIIc, IXa (red dash-dot line) and divisions VIIIa,b,d (red dashed line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

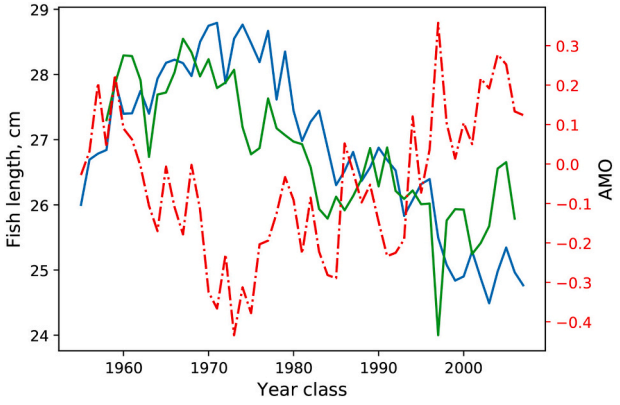


Fig. 6. The mean annual AMO index in the first year of life (red) and the mean length-at the 3 ring stage of herring in the CS (blue) and the NW (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the RY model was confirmed by rerunning the RY model ten times, using a different randomised test-train split each time and comparing the output. Neither the predicted relationships between the response and explanatory variables (as reflected in the partial dependence plots) nor the relative influence of the explanatory variables (as indicated by the variable influence plots) showed substantial variation between runs indicating that the model predictions were stable (Fig. 7a).

The MSE values of all the RI models were lower than those for the corresponding RY models (Table 3, Fig. 7). This confirms that splitting at the level of the response variable (RI models) can lead to over-estimation of the predictive capability of the model when explanatory variables are measured at a lower resolution than the response. The RY models provide a more robust test of the models' power by ensuring that the train and test sets contained unique combinations of explanatory variables.

The R squared values also showed that the models provided a better fit to the data for the CS population (RY model: $R^2 = 24.67$) compared

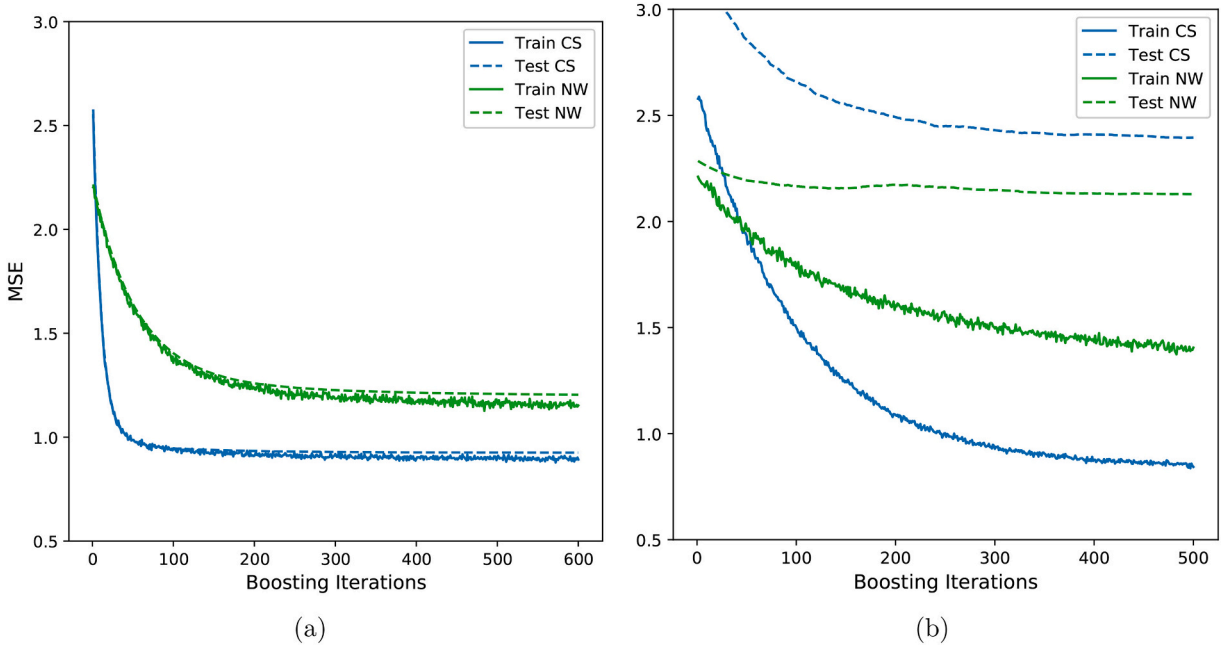


Fig. 7. Performance of GBRT models with a two-part split for the CS (green) and the NW (blue). The number of boosting iterations on the x-axis is plotted against MSE on the y-axis. In the RI model (a) the MSE curve for train data closely followed the MSE curve for the test data. In the RY model (b) there is a larger gap between train and test datasets, however train and test lines are still parallel to each other, which indicates stability of the model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
MSE's and R^2.

Model	splits	MSE	R^2
RI CS	2	0.92	66.25
RI CS	3	0.93	66.26
RI NW	2	1.20	46.11
RI NW	3	1.16	47.50
RY CS	2	2.39	24.67
RY NW	2	2.13	13.18

to the NW population (RY model: R^2 = 13.18). Overall, percentage explained variability was low. This reflects the high degree of individual variability in size that could not be accounted for by the explanatory variables, which capture the mean conditions experienced by the population.

3.4. Predicted trends in mean length

For the CS population, mean length of 3 ring herring predicted by the GBRT models followed similar temporal trends to observed mean length although the estimates from the RY model showed some divergence from the observed values, particularly in the mid 1970's and the 1990's (Fig. 3a). For the NW population the RY models tended to underestimate mean length early in the time series and overestimate it later in the time series (Fig. 3b).

3.5. Relative importance of the predictor variables

Fig. 8 displays the estimated relative importance of each variable in the GBRT models for the CS and the NW populations. The importance of the predictors varied between the RI and RY models, however, some general trends were apparent. Sea surface temperature (SST) in the first growing season was the most important predictor of length of 3 ring herring in both the CS and NW populations. The high variable influence score indicates that relative to the other variables, SST was selected most frequently for splitting and had the largest influence on predictive power.

For the CS population, the CPR estimated abundance of *Calanus finmarchicus* in area C3 was the second most important variable in both the RI and RY model although its influence much less marked than that of SST. Other descriptors of food availability (CPR estimated abundance of *Calanus finmarchicus* in area D4 and *Calanus helogalandicus* in areas C3 and D4) had only a minor influence on the model predictions for the CS population. Total population size made a small contribution to the predictions of both the RY and RI models while the importance of recruit abundance was more minor. Salinity, fishing pressure (mean life time fishing pressure - cumF and fishing pressure in year of capture - Fbar) and month of capture also had only a very minor influence on the predictions of both models.

For the NW population, variability in length was strongly associated with month of capture in both the RI and RY models, reflecting the expected increase in fish size as the year progresses. The next most important explanatory variable was mean lifetime fishing mortality. Density (Total N and recruitment), fishing pressure in year of capture, food supply (abundance of *Calanus finmarchicus* and *Calanus helagolandicus* in area C4) and salinity also made minor contributions to the model predictions of the NW models.

3.6. Partial dependence plots

The partial dependence plots display, for a selection of the more influential predictors, the marginal effects of each predictor on length of 3 ring herring. In general, the RI and RY models detect similar relationships between the predictors and the response although in some cases the RI model detected more complex relationships. This may reflect over-fitting of the data when splitting is implemented at the level of the individual (RI model) (Figs. 9a and Fig. 10a respectively). To avoid over emphasis on outlying data points, interpolations beyond the 9^{th} decile of the variable distribution were disregarded when interpreting the relationships between the predictors and response.

3.6.1. Sea surface temperature and salinity

In the CS population SSTs above 13.5$^{\circ}C$ were associated with a

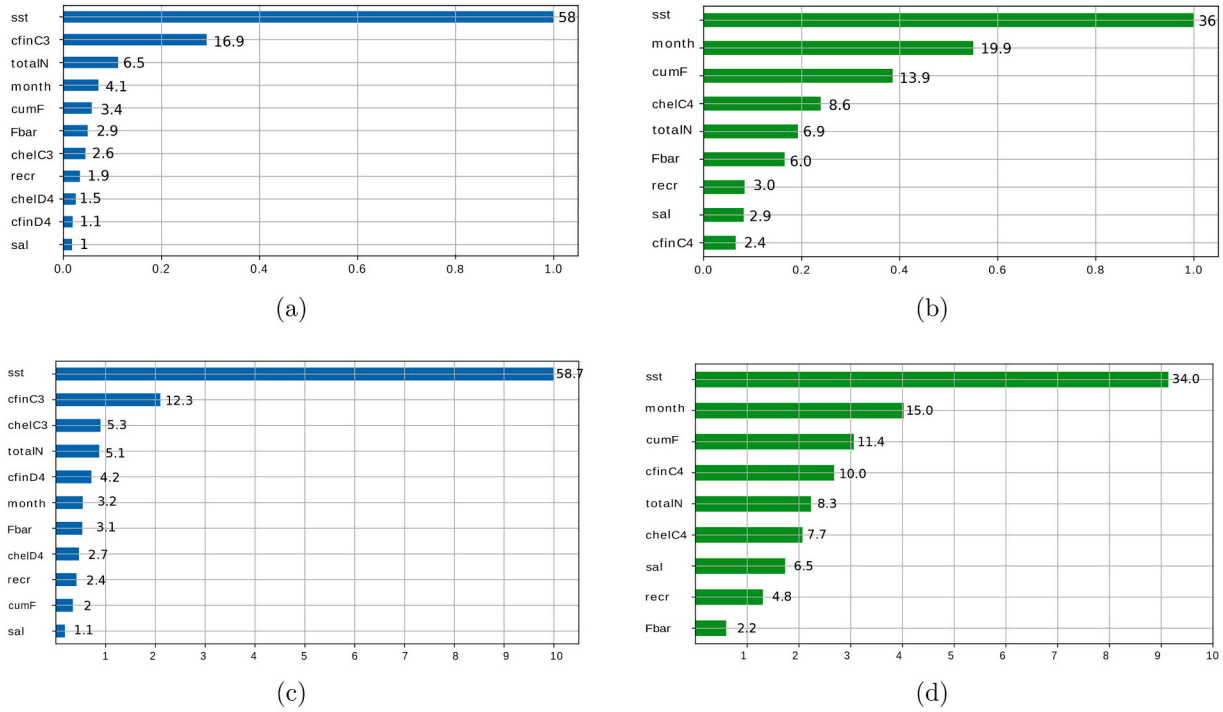


Fig. 8. Relative Variable Importance for the CS (a), (c) and the NW (b), (d) populations. RI models (a), and (b) were calculated from one GBRT model, whereas RY models (c), and (d) were based on 10 GBRT models. The VI scores as a proportion of the maximum for each model are shown on the x axis. The numbers beside each bar indicate the VI scores as a percentage of the total for each model. The numbers beside each bar indicate the VI scores as a percentage of the total for each model.

decrease in length. In the NW population the partial dependence plot showed a similar trend as in the Celtic Sea with a marked decrease in growth occurring around 12.9 − 13°C. For the NW population, length was predicted to increase with increasing salinity (RI and RY models), although the variable influence score indicated a weak effect (VI = 2.9 and 6.5 for the RI and RY models respectively.). Salinity had a negligible influence on length in the CS models (VI = 1 and 1.1 for the RI and RY models respectively).

3.6.2. Density dependence

For the CS population, the RI model predicted a decrease in length at higher population sizes which could be indicative of density dependence. However, according to the RY model the relationship between length and population size was highly non-linear with fish length decreasing at moderate population sizes and then increasing as population size increased. Similarly, the RI model predicted a decline in length of CS herring as recruitment increased from moderate (0.3×10^9 individuals) to high levels (1.3×10^9 individuals) while the RY model predicted an increase in length at high recruitment levels. For the NW population both the RI and the RY models predicted a decrease in length when population size increased from low to moderate levels followed by an increase at high density. Contrary to the expectation of density dependant growth, both the RY and RI models predicted a positive relationship between length and recruitment of NW herring. Overall, the analysis did not detect strong evidence of density dependant effects on growth in CS or NW herring.

3.6.3. Food availability

Relationships between length and food availability during the first growing season were variable and complex and often non-linear (Figs. 9 and 10). For the CS population the predominant effect detected by both the RI and the RY models was a positive relationship between length and the abundance of Calanus finmarchicus in area C3. Negative relationships with other food related variables (*Calanus helogalandicus* in area C3 and D4 and Calanus finmarchicus in area D4) were detected, but the variable influence scores indicated that these were of minor importance. For the NW population both the RI and the RY models predicted an increase in length at high abundance of *Calanus helogalandicus* in area C4, although the importance of this variable was lower for the RY model. The RY model predicted a negative relationship between length and the abundance of *Calanus finmarchicus* in area C4.

3.6.4. Fishing pressure

The strongest relationships with fishing pressure were observed in the NW population; the RY model predicted a decline in length as cumulative mean lifetime fishing pressure increased from 0.04 to 0.12 (predictions at higher fishing mortalities relied on less than 10% of the data and were not considered). The relationship predicted by the RI model was more complex and difficult to interpret. Other models predicted a positive relationship with Fbar but the Variable Importance scores (Fig. 8) suggested that the negative relationship with lifetime fishing pressure was the overriding effect.

3.6.5. Month of capture

Month of capture was included in the analysis to account for the increase in length as the season progressed. The expected increase in length was evident in the partial dependence plots for month of capture in the NW population. However, for the CS population the relationship was dome shaped. In this case April was treated as the first month of the year to align with the ageing convention for that stock. The partial dependence plot predicted a maximum length in October (month 7) which may reflect the dominance of the larger autumn spawning component in catches from that time of year.

3.6.6. Two way interactions between temperature and month

Two way partial dependence plots showed interactions between some pairs of explanatory variables. For the CS population, the dome shaped relationship between SST and month was most apparent at temperatures above 13.5°C; at lower temperatures neither the RI or the RY model predicted a decline in length later in the year (Fig. 10a). In the NW population the increase in length over the course of the year was most pronounced at temperatures above 13°C and commenced earlier in the year.

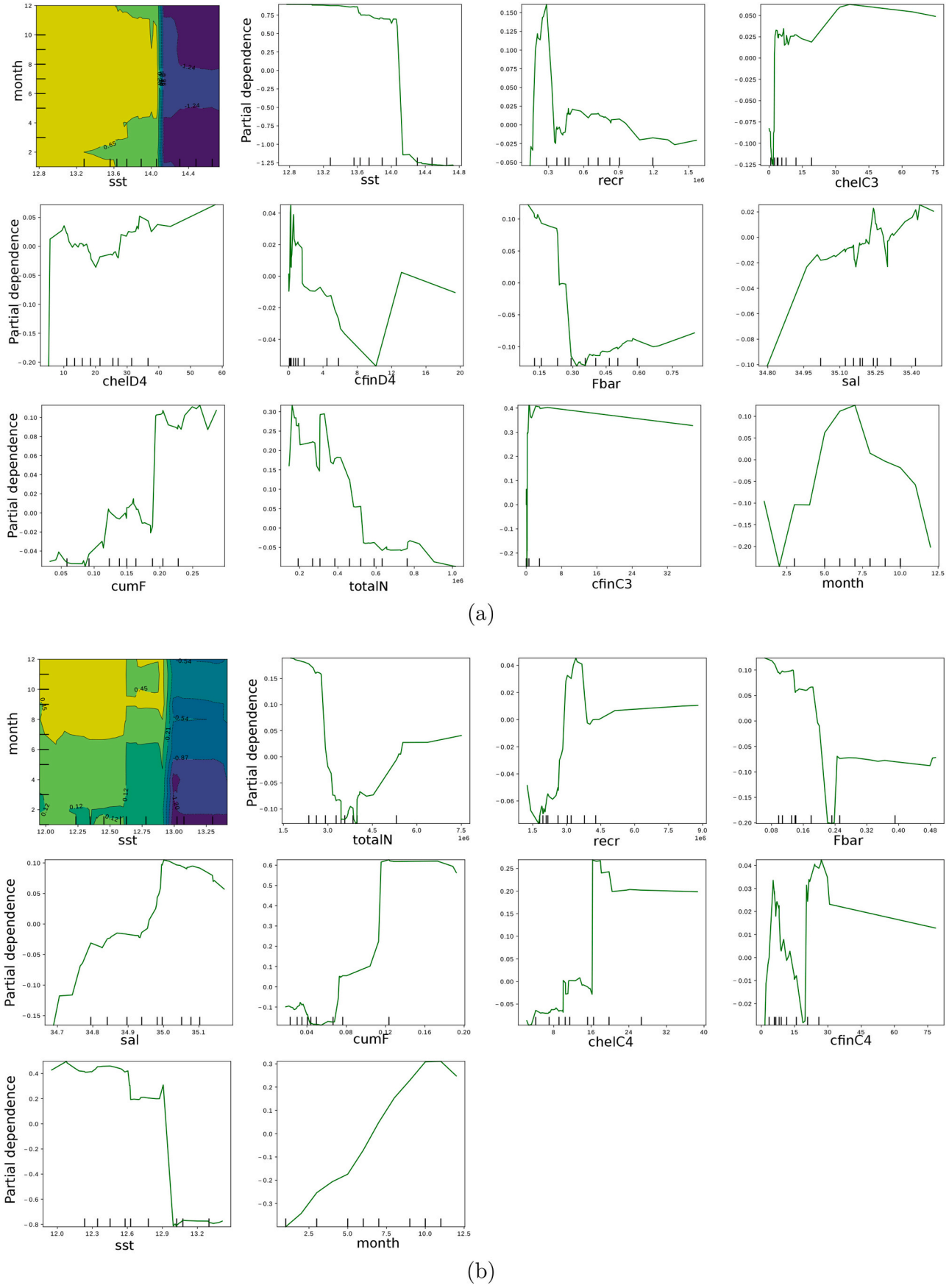
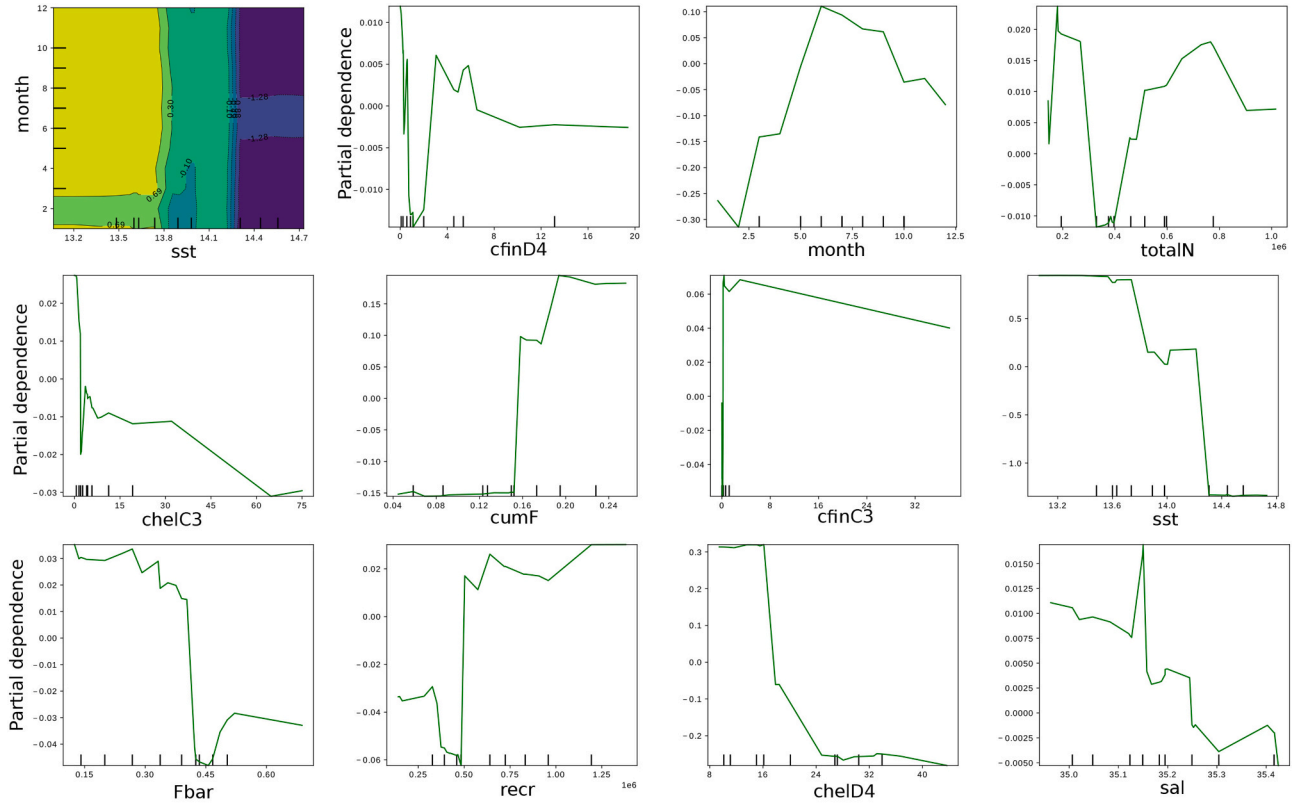
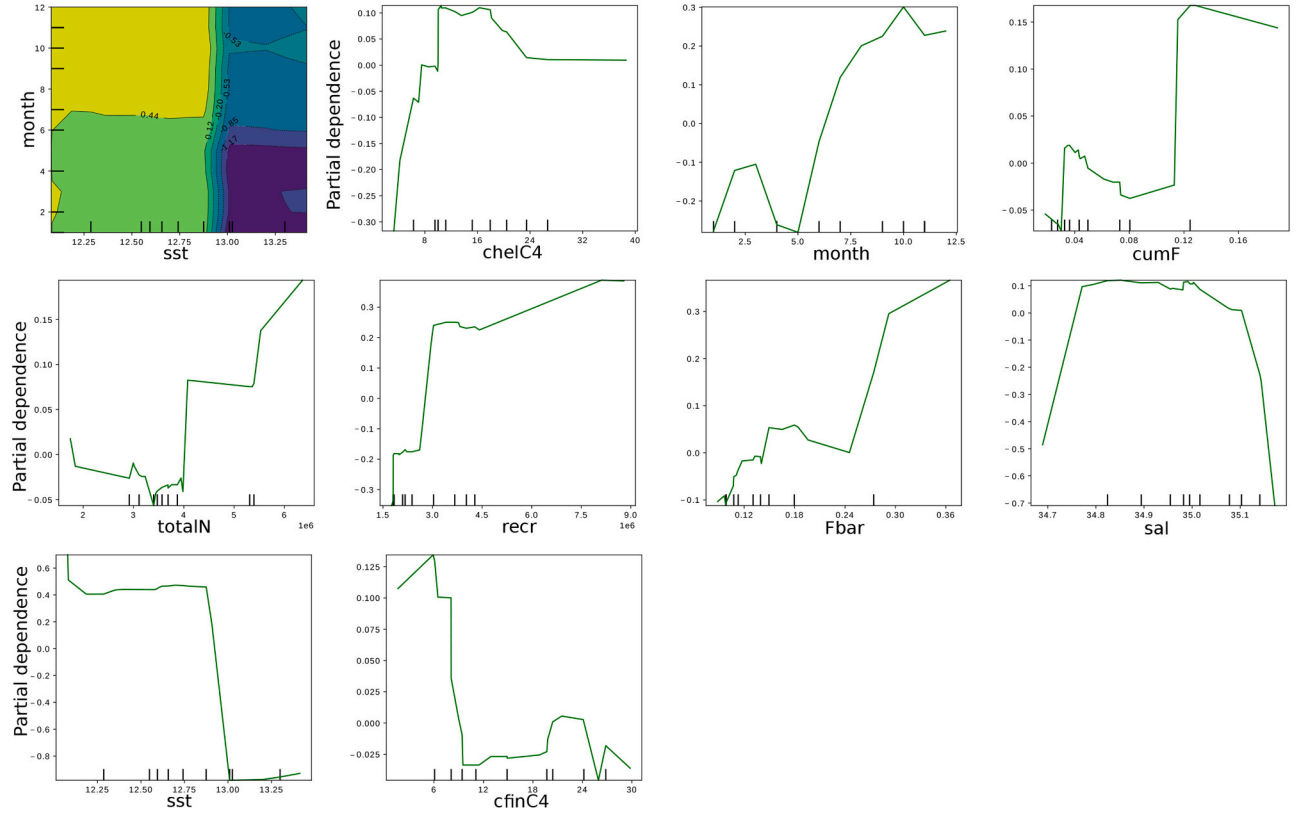


Fig. 9. RI partial dependence plots for the CS (a) and the NW (b). One-way partial dependence plots indicate the model predictions after the influence of a predictor marginalizing over all other predictors. Two-way partial dependence plots indicate interaction effect of varying degree. The colours are based on the contour levels and increase from purple to yellow. The extreme values for the axes are created using low (0.05) and high (0.95) percentiles. The hash marks on x-axis represent the deciles of the corresponding variable distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a)



(b)

Fig. 10. RY partial dependence plots for the CS (a) and the NW (b). One-way partial dependence plots indicate the model predictions after the influence of a predictor marginalizing over all other predictors. Two-way partial dependence plots indicate interaction effect of varying degree. The colours are based on the contour levels and increase from purple to yellow. The extreme values for the axes are created using low (0.05) and high (0.95) percentiles. The hash marks on x-axis represent the deciles of the corresponding variable distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

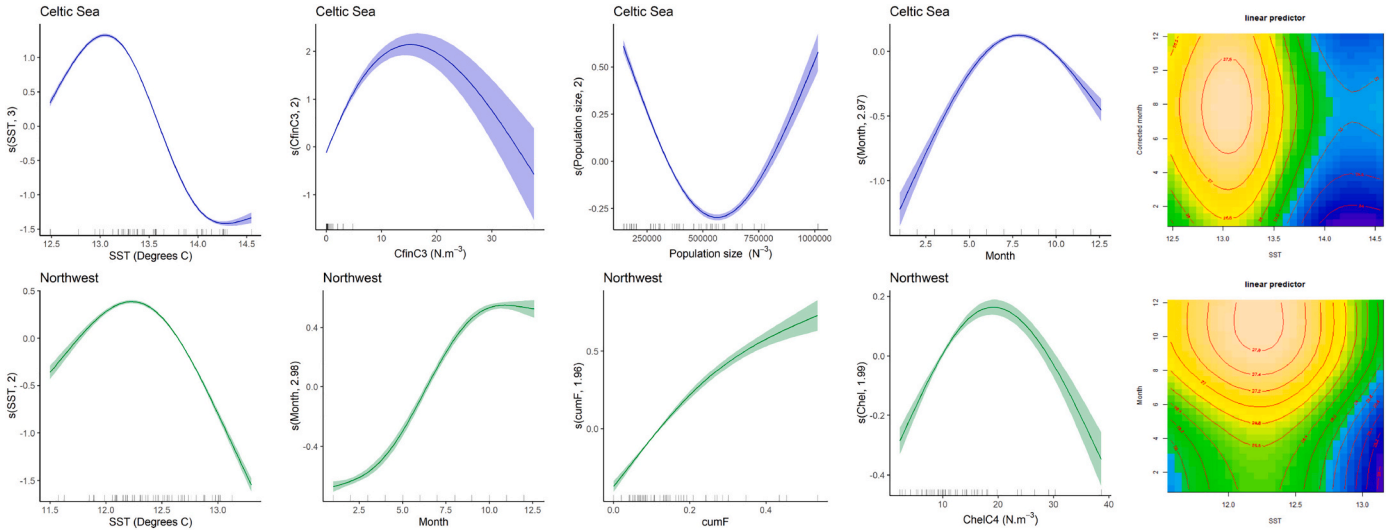


Fig. 11. GAM fits for the top four most influential variables, as indicated by the variable importance scores from the GBRT models. Two-way partial dependence plots indicate interaction effect of varying degree. The colours are based on the contour levels and increase from purple to yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.7. GAM modelling

GAM fits for the top four most influential variables were similar to those generated by the GBRT models for both populations (Fig. 11). The best fitting models included all four variables. The R^2 values were 39.9% and 26.7% for the CS and NW respectively. Two-way interactions between month and SST were also apparent in the GAM model fits, however the nature of the relationships were slightly different. In the CS, the decrease in size towards the end of the year was most apparent at around 13˚C and less evident at higher and lower temperatures. In the NW size increased earlier in the year at higher temperatures.

4. Discussion and conclusions

This study aimed to disentangle the effects of multiple drivers of variability on size-at-age of two herring populations across five decades and to identify the most likely causes of observed declines in growth Celtic Sea and North-West of Ireland herring since the 1980's. The advantages of a machine learning approach were combined with simple correlation analyses and ecological knowledge. The analysis demonstrated how GBRTs can be used to identify the relative importance of various exogenous variables in a dynamic system. GBRTs and GAMs detected similar relationships between length of 3 ring herring and key environmental and population related drivers and GBRT models accurately predicted the main trends in mean length in both populations. While both methods can be used to detect non-linear responses, an added advantage of the GBRT approach is that the relationships between response and predictor variables and interactions between predictors do not need to be explicitly specified in the model (Leathwick et al., 2006). A potential drawback of GBRTs is that if an entirely data-driven approach is applied to the analysis, spurious relationships with no biological or ecological basis may be detected entirely by chance (Elliot et al. 2016). This was avoided in our analysis by selecting the predictor variables based on knowledge of the species' ecology and observed correlations with climatic indices. In addition, the robustness of the analysis was ensured by testing model predictions using combinations of predictor variables that were held back during the model testing stage (the RY models) and by splitting the data into three sets (training, validation and testing) to evaluate performance of the RI model. That the detected relationships with the predictor variables and their relative importance were remarkably consistent across all of the approaches tested inspires confidence in the results. The relationships detected by the GBRT models are complex and interactive and do not necessarily indicate causation. However, many of the observed relationships have a plausible ecological basis. The comparison of the Celtic Sea and North West herring populations shows both common responses to global drivers and more localised stock-specific relationships. In particular, a marked decrease in length above a threshold temperature was detected in both populations.

In the CS and the NW populations, length of 3 ring herring was negatively associated with mean AMO in the year after hatching. The observed negative correlation remained significant after correction for temporal auto-correlation and is consistent with previous reports of climate driven multi-decadal fluctuations in the distribution and abundance of small pelagic clupeids (Alheit et al., 2014). During a previous warm period, that coincided with a positive phase of the AMO (1930's–1960's), the abundance of herring in the English Channel (at the southern limits of the species distribution) decreased and its distribution contracted (Southward et al., 1988) while abundance of the more northerly Norwegian spring spawning herring population increased (Toresen and Østvedt, 2000). Concurrent changes in the dynamics of other pelagic clupeids (anchovy, sardine, sardinella and sprat) and similar trends during the more recent warming period (after the mid-1990's) are indicative of climate driven ecosystem regime shifts that may be reflected in the AMO signal (Alheit et al., 2014; Edwards et al., 2013). The results of this study demonstrate that climatic fluctuations may manifest as changes in fish growth as well as abundance. The association between fish size and AMO was stronger and more consistent in the Celtic Sea than in the Northwest; possibly indicating that the Celtic Sea population, existing close to the southern limit of the species' distribution, is thermally limited and more vulnerable to climatic warming than the more northerly Northwest herring population.

Fluctuations in AMO represent broad-scale environmental change that can affect multiple trophic levels with direct and indirect consequences for fisheries. While the AMO is primarily an index of SST anomalies, it is also associated with regional fluctuations in precipitation, wind patterns, ocean circulation wind mixing and stratification (Nye et al., 2014). AMO related increases in SST can intensify the impact of climate change and directly impact on fish survival, growth and other physiological processes (Nye et al., 2009). The combined effects of the AMO on temperature, wind patterns and stratification are thought to drive decadal changes in global phytoplankton abundance (Martinez et al., 2009; Nye et al., 2014), while temperature changes

lead to changes in the distribution of zooplankton, with consequences for growth of planktivorous fish (Beaugrand et al., 2009; McGinty et al., 2011). Several interacting mechanisms may therefore underlie observed synchronicities between the AMO signal and pelagic fish populations.

Here, the GBRT modelling approach provided a means to identify the variables that most likely underlie the observed association between the AMO index and herring length-at-age. The low to moderate R squared values associated with GBRT models (13.2–66.3%) and the GAMs (26.7–39.9%) reflected the fact that individual variability in length-at-age could not be accounted for by the explanatory variables which were measured with broad temporal and spatial resolution and with a degree of uncertainty that is not accounted for in the model. The herring time series extended back to the 1950's; highly spatially resolved data describing the environmental and biological drivers of growth are not available for this time period. Nonetheless, predicted mean-lengths from the GBRT models were close to the observed mean lengths for most of the time series (Fig. 3) and the observed decline in mean length of 3 ring herring in the Celtic Sea population after 1980 was evident in the predictions from both the RY and RI models. The R squared values also showed that model predictions were more accurate for the Celtic Sea population than the Northwest population. This was not due to levels of individual variability in length-at-age which were similar for the two stocks. It may be that in the warmer waters of the Celtic Sea increasing temperatures have had a stronger limiting influence on growth of herring, while in the Northwest, other local drivers, not included in our analysis, might make a larger contribution to variation in length-at-age.

The predominant signal detected by all of the GBRT models was a negative association between length of 3 ring herring and SST during the first growing season (April–August). Both the GBRT models and the GAMs predicted a non-linear response to temperature; an increase in length at lower temperatures followed by a steep decline at the upper end of the temperature range was evident from the GAM fits while the GBRT detected a step-change in length above a threshold temperature. As expected, length of 3 ring herring was related to month of capture; in the NW population length increased throughout the year whereas in the Celtic Sea length increased to a maximum in October and then decreased, probably due to the increased contribution of the smaller sized winter spawning stock component to catches later in the year (Harma et al., 2012). The two way partial dependence plots from the GBRT and the GAM models indicated that the relationships between length and month of capture varied with temperature. The interaction between seasonal growth and temperature was complex and the detected relationships varied between modelling approaches. The most consistent effect was an earlier increase in size of NW herring during warmer years which may indicate that the timing of seasonal growth is changing with increasing temperatures.

At northern latitudes, growth of juvenile herring tends to increase with increasing temperature (Husebø et al., 2007). However, an experimental study conducted using herring from the south eastern North Sea recorded metabolic optimum at around 15–16.1°C with a subsequent decrease in metabolic rate occurring with further increases in temperature (Bernreuther et al., 2013). Water temperatures in the Celtic Sea and the Northwest are cooler than the south-eastern North Sea and the thermal optimum for these populations may therefore be lower due to genetic influences on growth and metabolism. Mean SST in August ranged from 15.1°C to 18°C in the Celtic Sea and from 13.4°C to 15.7°C in the Northwest. As sea temperatures rose due to the combined effect of a positive phase of the AMO and climate change, herring were more likely to encounter temperatures that were sub-optimal for growth and metabolism, particularly in the Celtic Sea. In the absence of individual temperature histories, it is not possible to determine if this exposure would be sufficiently frequent or prolonged to produce such a pronounced decline in growth as has been observed in Celtic Sea herring. However, it is plausible that the direct effects of increasing temperature on growth and metabolism could at least partly contribute to the decline.

The results highlight how the response of a population to increasing temperature can vary to what is predicted for the species as a whole. In a cross-population examination of weight-length relationships, Brunel and Dickey-Collas (2010) observed that growth rate of herring was positively correlated with mean annual SST while asymptotic weight was negatively correlated. The authors proposed that global warming could lead to higher growth of young age classes, slower growth of older individuals and a shorter lifespan of herring. However, temperature-growth relationships within populations were difficult to resolve; both negative and positive correlations were detected but many were non-significant. In Pacific herring (*Clupea pallasai*) a positive correlation between growth during the first year and SST broke down at higher temperatures and marked reductions in growth were observed in years when mean July–August temperatures exceeded 12.5°C (Batten et al., 2016). In addition, Watanabe et al. (2008) showed that high temperatures during winter had a negative effect on growth of Hokkaido spring spawning Pacific herring which occupy the southern boundary of the distribution range for that species. In Atlantic herring populations, predicted positive temperature-growth relationships might not apply when temperatures exceed a certain population-specific threshold. This signals caution when extrapolating climate change effects from contemporary field observations and highlights the importance of considering biological responses at distributional extremes.

Changes in the growth of planktivorous fish populations can be mediated by trophic interactions via the influence of environmental processes on plankton abundance (bottom-up control) or due to changes in population density and food-competition (top-down control). Changes in size-at-age have been attributed to density dependence in Baltic Sea (Cardinale and Arrhenius, 2000) and George's Bank (Melvin and Stephenson, 2007) herring populations. In the North Sea, the influence of hydrographic conditions on plankton abundance drives short-term interannual variability in growth of larval and juvenile herring while long-term trends are driven by density dependence (Dickey-Collas et al., 2010; Heath et al., 1997; Shin and Rochet, 1998). In the present study, the abundance of some prey items during the first growing season was negatively correlated with length of 3 ring herring in both populations while relationships with recruitment strength were positive (NW population) or weakly negative (CS population). This is indicative of bottom-up control of food supply rather than density dependence. A positive correlation between length and recruitment strength could arise if conditions which favour high survival during the first year are also favourable for subsequent growth. For both populations, negative relationships between length and adult population size in the year of capture were detected, indicating that adult growth is subject to density dependence. Overall, associations with food availability and density were detected but these were weak and of minor explanatory importance relative to the temperature effect. The results suggest that food availability and density do not exert a strong influence on temporal trends in length of Celtic Sea and Northwest herring.

While length and weight at age declined markedly in Celtic Sea herring after the 1980's there was no concomitant decline in fish condition index in either the Celtic Sea or Northwest population (mean weight at 27 cm: Fig. 4). This provides further indication that pronounced changes in feeding conditions have not occurred. This is in contrast to the situation in the Baltic Sea; declines in weight-at-age of herring during the 1990's coincided with a decrease in condition and both were attributed to increasing densities of pelagic fish and consequent reductions in food availability (Cardinale and Arrhenius, 2000). Long-term studies have reported climate related shifts in the distribution of zooplankton species in the Northeast Atlantic with northward movement of temperature species such as *Calanus helgolandicus* and declines in boreal species like *Calanus finmarchicus* (Pitois and Fox, 2006). However, the abundance of these calanoid copepods in the Celtic Sea has remained relatively unchanged (Lauria et al., 2012;

Pitois and Fox, 2006). In addition, from 1986 to 2007, temporal trends in food availability were not linked to biomass of juvenile herring or to climate indicator (Lauria et al., 2012). It appears unlikely that the availability of calanoid copepods is a major driver of observed declines in growth of Celtic Sea herring. Nonetheless, it must be borne in mind that the CPR provides a spatially coarse index of plankton abundance and may not capture localised fluctuations in food availability, particularly within coastal areas inhabited by young feeding herring.

Intensive fishing is predicted to alter the growth rate of fish populations directly, by selective removal of large and fast growing individuals from the population, or indirectly through selection for earlier maturation (Heino et al., 2015). Reductions in size-at-age due to the selective effects of fishing have been demonstrated in Atlantic cod *Gadus morhua* in the Gulf of St Lawrence (Swain et al., 2007). In the present study, mean lifetime fishing mortality contributed very little to variability in length of 3 ring herring in the Celtic Sea herring but was a relatively important explanatory variable in the models describing size variation in the Northwest population. For the NW population there was some evidence of a decline in length with increasing mean lifetime fishing pressure. For the CS population the detected relationship between length and mean lifetime fishing pressure was positive. It is likely that this association indicates temporal concurrence rather than causation; fishing effort was restricted during the 1980's and 2000's in response to declines in stock biomass which coincided with declines in size-at-age (ICES, 2016). Overall, the analysis did not find strong evidence of fishing induced changes in size in the CS or NW herring populations although the abundance of populations has undeniably been affected by decades of intensive fishing pressure.

Due to the nature of the herring fisheries, evolutionary responses to fishing may be relatively minor (Engelhard and Heino, 2004). In the waters around Ireland, herring are targeted predominantly by midwater pair trawls, primarily during the spawning season. Trawls are known to capture a wide range of length classes (Kuparinen et al., 2009). Length-dependant escapement of Baltic Sea herring from commercial trawls has been demonstrated for fish below 15 cm length (Suuronen et al., 1997). However, escaped fish suffer high rates of mortality (Suuronen et al., 1996). Herring fisheries target aggregations of spawning or feeding adults (> 19cm) which tend to be spatially segregated from immature fish (Clarke et al., 2010; O'Donnell et al., 2015). It is plausible that probability of capture for mature herring on the spawning grounds is not strongly size dependant. However, experimental investigation of gear selectivity at the fishing grounds would be needed to confirm this. The possible contribution of fishing to declines in size-at-age of Celtic Sea herring could be further interrogated using the approach of Swain et al. (2007) who related back-calculated lengths-at-age to cohort specific estimates of fisheries induced selection while controlling for temperature and density related changes in growth.

It is possible that the decline in growth of herring in the Celtic Sea is symptomatic of broad-scale changes in the pelagic ecosystem. Clupeid fisheries are characterised by variable abundance. In the past, cyclical changes in herring populations have coincided with alternate changes in sardine (*Sardina pilchardus*) populations, apparently in response to climatic fluctuations and associated ecosystem change (Alheit et al., 2014; Southward et al., 1988). Although sardines occur in the Celtic Sea and are commercially exploited, the stock in this area is not assessed and there is no biological data available from scientific surveys or the commercial catch (ICES, 2017; Marine Institute, 2016). Available data for adjacent areas do indicate recent changes in the growth of sardine. Trends in weight-age indicate that growth rates of sardine in the southern Bay of Biscay and Iberian coast have increased since 1990, while growth rates in the wider Bay of Biscay area have decreased since 2000. Without extended growth time-series, it is not possible to establish if these trends reflect a synchronous response to climatic drivers. However, the patterns indicate that closer monitoring of the distribution, growth and population dynamics of pelagic clupeids in the Celtic Sea is warranted, particularly as new fisheries develop.

In summary, the results of the analysis indicate that the declines in size-at-age of Celtic Sea herring since the 1980's are most strongly associated with increasing sea temperatures as a consequence of climate change and with a positive AMO index. The more stable size-at-age trends in Northwest herring may reflect the stock's more northerly distribution and lower exposure to metabolically sub-optimal temperatures, although that population also displays temperature related declines in size which indicate a vulnerability to future temperature rises.

Analysis was performed:

```
Platform version: Linux-4.4.0-31-generic-x86i_64-
with-debian-stretch-sid
Python 3.6.5 :: Anaconda, Inc.
[GCC 4.4.7 20120313 (Red Hat 4.4.7-1)]
Pandas version: 0.24.2
Matplotlib version: 3.0.2
sklearn version 0.20.2
```

Declaration of Competing Interest

None.

Acknowledgements

This research was supported by a grant from the Irish Environmental Protection Agency (Ecosystem tipping points: learning from the past to manage for the future, project code 2015-NC-MS-3). CH received a PhD scholarship under the Sea Change strategy supported by the Marine Institute and the Marine Research Sub-Program of the National Development Plan 2007–2013 (Grant-Aid Agreement: PhD/FS/07/007 (01))

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2020.101154.

References

Alheit, J., Licandro, P., Coombs, S., Garcia, A., Giráldez, A., Santamaría, M.T.G., Slotte, A., Tsikliras, A., 2014. Atlantic Multidecadal Oscillation (AMO) modulates dynamics of small pelagic fishes and ecosystem regime shifts in the eastern North and Central Atlantic. J. Mar. Syst. 131, 21–35.

Audzijonyte, A., Fulton, E., Haddon, M., Helidoniotis, F., Hobday, A.J., Kuparinen, A., Morrongiello, J., Smith, A.D.M., Upston, J., Waples, R.S., 2016. Trends and management implications of human-influenced life-history changes in marine ectotherms. Fish Fish. 17, 1005–1028.

Batten, S.D., Moffitt, S., Pegau, W.S., Campbell, R., 2016. Plankton indices explain interannual variability in Prince William Sound herring first year growth. Fish Oceonagr.. 25, 420–432.

Baudron, A.R., Needle, C.L., Marshall, C.T., 2011. Implications of a warming north sea for the growth of haddock melanogrammus aeglefinus. J. Fish Biol. 78, 1874–1889.

Beaugrand, G., Luczak, C., Edwards, M., 2009. Rapid biogeographical plankton shifts in the North Atlantic Ocean. Glob. Chang. Biol. 15, 1790–1803.

Berg, F., Almeland, O.W., Skadal, J., Slotte, A., Andersson, L., Folkvord, A., 2018. Genetic factors have a major effect on growth, number of vertebrae and otolith shape in Atlantic herring (*Clupea harengus*). PLoS One 13, 1–16.

Bernreuther, M., Herrmann, J., Peck, M.A., Temming, A., 2013. Growth energetics of juvenile herring, *Clupea harengus L.*: food conversion efficiency and temperature dependency of metabolic rate. J. Appl. Ichthyol. 29, 331–340.

Brander, K.M., 2007. The role of growth changes in the decline and recovery of North Atlantic cod stocks since 1970. ISEC J. Mar. Sci. 64, 211–217.

Brophy, D., Danilowicz, B.S., 2003. The influence of pre-recruitment growth on subsequent growth and age at first spawning in Atlantic herring (*Clupea harengus* L.). ICES J. Mar. Sci. 60, 1103–1113.

Brunel, T., Dickey-Collas, M., 2010. Effects of temperature and population density on von Bertalanffy growth parameters in Atlantic herring: a macro-ecological analysis. Mar. Ecol. Prog. Ser. 405, 15–28.

Cameron, M., Lucieer, V., Barrett, N., Johnson, C., Edgar, G., 2014. Understanding community-habitat associations of temperate reef fishes using fine-resolution bathymetric measures of physical structure. Mar. Ecol. Prog. Ser. 506, 213–229.

Cardinale, M., Arrhenius, F., 2000. Decreasing weight-at-age of Atlantic herring (*Clupea*

harengus) from the Baltic Sea between 1986 and 1996: a statistical analysis. ICES J. Mar. Sci. 57, 882–893.
Casini, M., Cardinale, M., Hjelm, J., 2006. Inter-annual variation in herring, *Clupea harengus,* and sprat, *Sprattus sprattus,* condition in the Central Baltic Sea: what gives the tune? Oikos 112, 638–650.
Chelton, D.B., 1984. Commentary: short-term climatic variability in the Northeast Pacific Ocean. In: Pearcy, W. (Ed.), The Influence of Ocean Conditions on the Production of Salmonids in the North Pacific. Oregon State University Press, Corvallis, Oreg., pp. 87–99.
Clarke, M., Egan, A., Molloy, J., 2010. A Survey of Nursery Grounds for Celtic Sea and VIIj Herring. Report.
Conover, D.O., Munch, S.B., 2002. Sustaining fisheries yields over evolutionary time scales. Science 297, 94–96.
Crain, C.M., Kroeker, K., Halpern, B.S., 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecol. Lett. 11, 1304–1315.
Dickey-Collas, M., Nash, R.D.M., Brunel, T., van Damme, C.J.G., Marshall, C.T., Payne, M.R., Corten, A., Geffen, A.J., Peck, M.A., Hatfield, E.M.C., Hintzen, N.T., Enberg, K., Kell, L.T., Simmonds, E.J., 2010. Lessons learned from stock collapse and recovery of North Sea herring: a review. ICES J. Mar. Sci. 67, 1875–1886.
Edwards, M., Beaugrand, G., Helaouët, P., Alheit, J., Coombs, S., 2013. Marine ecosystem response to the Atlantic Multidecadal Oscillation. PLoS One 8, 1–5.
Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813.
Elliott, K.C., Cheruvelil, K.S., Montgomery, G.M., Soranno, P.A., 2016. Conceptions of good science in our data-rich world. BioScience 66, 880–889.
Enfield, D.B., Mestas-Nunez, A.M., Trimble, P.J., 2001. The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. Geophys. Res. Lett. 28, 2077–2080.
Engelhard, G.H., Heino, M., 2004. Maturity changes in Norwegian spring-spawning herring *Clupea harengus*: compensatory or evolutionary responses? Mar. Ecol. Prog. Ser. 272, 245–256.
Escobar-Flores, P., O'Driscoll, R., Montgomery, J., 2013. Acoustic characterization of pelagic fish distribution across the South Pacific Ocean. Mar. Ecol. Prog. Ser. 490, 169–183.
Franklin, E., Jokiel, P., Donahue, M., 2013. Predictive modeling of coral distribution and abundance in the Hawaiian Islands. Mar. Ecol. Prog. Ser. 481, 121–132.
Friedman, J.H., 2000. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.
Froeschke, J.T., Froeschke, B.F., 2016. Two-stage boosted regression tree model to characterize southern flounder distribution in Texas estuaries at varying population sizes. Mar Coast Fish 8, 222–231.
Griffen, B.D., Belgrad, B.A., Cannizzo, Z.J., Knotts, E.R., Hancock, E.R., 2016. Rethinking our approach to multiple stressor studies in marine environments. Mar. Ecol. Prog. Ser. 543, 273–281.
Harma, C., Brophy, D., Minto, C., Clarke, M., 2012. The rise and fall of autumn-spawning herring (*Clupea harengus* L.) in the Celtic Sea between 1959 and 2009: temporal trends in spawning component diversity. Fish. Res. 121–122, 31–42.
Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning. Springer.
Heath, M., Scott, B., Bryant, A., 1997. Modelling the growth of herring from four different stocks in the North Sea. J. Sea Res. 38, 413–436.
Heino, M., Pauli, B.D., Dieckmann, U., 2015. Fisheries-induced evolution. Annu. Rev. Ecol. Evol. Syst. 46, 461–480.
Huse, G., Toresen, R., 1996. A comparative study of the feeding habits of herring (*Clupea harengus, Clupeidae,* 1.) and capelin (*Mallotus villosus, Osmeridae, Müller*) in the Barents Sea. Sarsia 81, 143–153.
Husebø, A., Slotte, A., Stenevik, E.K., 2007. Growth of juvenile Norwegian spring-spawning herring in relation to latitudinal and interannual differences in temperature and fish density in their coastal and fjord nursery areas. ICES J. Mar. Sci. 64, 1161–1172.
ICES, 2014. Report of the Herring Assessment Working Group for the Area South of 62N (HAWG), 11-20 March 2014. ICES HQ, Copenhagen, Denmark ICES CM 2014/ACOM:06. 1257 pp. Report.
ICES, 2015. Report of the Herring Assessment Working Group for the Area South of 62N (HAWG), 10-19 March 2015. ICES HQ, Copenhagen, Denmark ICES CM 2015/ACOM:06. 850 pp. Report.
ICES, 2016. Report of the Herring Assessment Working Group for the Area South of 62N (HAWG), 29 March-7 April 2016. ICES HQ, Copenhagen, Denmark ICES CM 2016/ACOM:07. 867 pp. Report.
ICES, 2017. Report of the Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA) 24-29 June 2017. Bilbao, Spain. ICES CM 2017/ACOM:17. 602 pp. Report.
Ingleby, B., Huddleston, M., 2007. Quality control of ocean temperature and salinity profiles – historical and real-time data. J. Mar. Syst. 65, 158–175.
ISEC, 2019. Report of the Herring Assessment Working Group for the Area South of 62N (HAWG). 1:2. 971 pp. https://doi.org/10.17895/ices.pub.5460. Report.
Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., Hooker, G., 2009. Data-intensive science: a new paradigm for biodiversity studies. BioScience 59, 613–620.
Knight, J.R., Folland, C.K., Scaife, A.A., 2006. Climate impacts of the Atlantic Multidecadal Oscillation. Geophys. Res. Lett. 33.
Koenigstein, S., Mark, F.C., Gossling-Reisemann, S., Reuter, H., Poertner, H.O., 2016. Modelling climate change impacts on marine fish populations: process-based integration of ocean warming, acidification and other environmental drivers. Fish Fish. 17, 972–1004.
Kuparinen, A., Kuikka, S., Merilä, J., 2009. Estimating fisheries-induced selection: traditional gear selectivity research meets fisheries-induced evolution. Evol. Appl. 2, 234–243.
Lauria, V., Attrill, M.J., Pinnegar, J.K., Brown, A., Edwards, M., Votier, S.C., 2012. Influence of climate change and trophic coupling across four trophic levels in the Celtic Sea. PLoS One 7, e47208.
Law, R., 2000. Fishing, selection, and phenotypic evolution. ICES J. Mar. Sci. 57, 659–668.
Leathwick, J., Elith, J., Francis, M., Hastie, T., Taylor, P., 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Mar. Ecol. Prog. Ser. 321, 267–281.
Lehodey, P., Alheit, J., Barange, M., Baumgartner, T., Beaugrand, G., Drinkwater, K., Fromentin, J.M., Hare, S.R., Ottersen, G., Perry, R.I., Roy, C., van der Lingen, C.D., Werner, F., 2006. Climate variability, fish, and fisheries. J. Clim. 19, 5009–5030.
Lynch, D., 2011. Biological Changes in Celtic Sea and Southwest of Ireland Herring, Based on a Long-Term Data Archival Project. Master's thesis. Trinity College Dublin, Ireland.
Maloney, K.O., Schmid, M., Weller, D.E., 2012. Applying additive modelling and gradient boosting to assess the effects of watershed and reach characteristics on riverine assemblages. Methods Ecol. Evol. 3, 116–128.
Marine Institute, 2016. The Stock Book 2016: Annual Review of Fish Stocks in 2016 with Management Advice for 2017. Report. Marine Institute.
Martinez, E., Antoine, D., D'Ortenzio, F., Gentili, B., 2009. Climate-driven basin-scale decadal oscillations of oceanic phytoplankton. Science 326, 1253–1256.
McGinty, N., Power, A., Johnson, M., 2011. Variation among Northeast Atlantic regions in the responses of zooplankton to climate change: not all areas follow the same path. J. Exp. Mar. Biol. Ecol. 400, 120–131.
Melvin, G.D., Stephenson, R.L., 2007. The dynamics of a recovering fish stock: Georges Bank herring. ICES J. Mar. Sci. 64, 69–82.
Muttil, N., Chau, K.W., 2007. Machine-learning paradigms for selecting ecologically significant input variables. Eng. Appl. Artif. Intell. 20, 735–744.
Neuheimer, A.B., Taggart, C.T., 2010. Can changes in length-at-age and maturation timing in Scotian shelf haddock (*Melanogrammus aeglefinus*) be explained by fishing? Can. J. Fish. Aquat. Sci. 67, 854–865.
Neuheimer, A.B., Thresher, R.E., Lyle, J.M., Semmens, J.M., 2011. Tolerance limit for fish growth exceeded by warming waters. Nat. Clim. Chang. 1, 110–113.
Nye, J.A., Link, J.S., Hare, J.A., Overholtz, W.J., 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Mar. Ecol. Prog. Ser. 393, 11–129.
Nye, J.A., Baker, M.R., Bell, R., Kenny, A., Kilbourne, K.H., Friedland, K.D., Martino, E., Stachura, M.M., Houtan, K.S.V., Wood, R., 2014. Ecosystem effects of the Atlantic Multidecadal Oscillation. J. Mar. Syst. 133, 103–116.
O'Donnell, C., Lynch, D., Lyons, K., Keogh, N., O'Donovan, M., 2015. Celtic Sea Herring Acoustic Survey Cruise Report. Report. Marine Institute FSS Survey Series: 2015/04.
Olden, J.D., Lawler, J.J., Poff, N.L., 2008. Machine learning methods without tears: a primer for ecologists. Q. Rev. Biol. 83, 171–193.
Ottersen, G., Loeng, H., 2000. Covariability in early growth and year-class strength of Barents Sea cod, haddock, and herring: the environmental link. ICES J. Mar. Sci. 57, 339–348.
Pedersen, T., Fossheim, M., 2008. Diet of 0-group stages of capelin (*Mallotus villosus*), herring (*Clupea harengus*) and cod (*Gadus morhua*) during spring and summer in the Barents Sea. Mar. Biol. 153, 1037–1046.
Perry, R.I., Cury, P., Brander, K., Jennings, S., Mollmann, C., Planque, B., 2010. Sensitivity of marine systems to climate and fishing: concepts, issues and management responses. J. Mar. Syst. 79, 427–435.
Peters, D.P.C., Havstad, K.M., Cushing, J., Tweedie, C., Fuentes, O., Villanueva-Rosales, N., 2014. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. Ecosphere 5, 1–15.
Pitois, S.G., Fox, C.J., 2006. Long-term changes in zooplankton biomass concentration and mean size over the Northwest European shelf inferred from Continuous Plankton Recorder data. ICES J. Mar. Sci. 63, 785–798.
Pittman, S.J., Costa, B.M., Battista, T.A., 2009. Using lidar bathymetry and boosted regression trees to predict the diversity and abundance of fish and corals. J. Coast Res. 27–38.
Pyper, B.J., Peterman, R.M., 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. Can. J. Fish. Aquat. Sci. 55, 2127–2140.
Rajasilta, M., Laine, P., Paranko, J., 2011. Current growth, fat reserves and somatic condition of juvenile Baltic herring (*Clupea harengus membras*) reared in different salinities. Helgol. Mar. Res. 65, 59–66.
Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily high-resolution-blended analyses for sea surface temperature. J. Clim. 20, 5473–5496.
Richardson, A., Walne, A., John, A., Jonas, T., Lindley, J., Sims, D., Stevens, D., Witte, M., 2006. Using continuous plankton recorder data. Prog. Oceanogr. 68, 27–74.
Rogers, J., 1997. North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe. J. Clim. 10, 1635–1647.
Shin, Y.J., Rochet, M.J., 1998. A model for the phenotypic plasticity of North Sea herring growth in relation to trophic conditions. Aquat. Living Resour. 11, 315–324.
Southward, A.J., Boalch, G.T., Maddock, L., 1988. Fluctuations in the herring and pilchard fisheries of Devon and Cornwall linked to change in climate since the 16th century. J. Mar. Biol. Assoc. UK 68, 423–445.
Sugihara, G., May, R.M., 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344, 734–741.
Suuronen, P., Erickson, D.L., Orrensalo, A., 1996. Mortality of herring escaping from pelagic trawl codends. Fish. Res. 25, 305–321.
Suuronen, P., Lehtonen, E., Wallace, J., 1997. Avoidance and escape behaviour by herring encountering midwater trawls. Fish. Res. 29, 13–24.

Swain, D.P., Sinclair, A.F., Mark Hanson, J., 2007. Evolutionary response to size-selective mortality in an exploited fish population. Proc. R. Soc. B 274, 1015–1022.
Toresen, R., Østvedt, O., 2000. Variation in abundance of Norwegian spring-spawning herring (*Clupea harengus, Clupeidae*) throughout the 20th century and the influence of climatic fluctuations. Fish Fish. 1, 231–256.
Trigal, C., Degerman, E., 2015. Multiple factors and thresholds explaining fish species distributions in lowland streams. Glob. Ecol. Conserv. 4, 589–601.
van Walraven, L., Mollet, F.M., van Damme, C.J.G., Rijnsdorp, A.D., 2010. Fisheries-induced evolution in growth, maturation and reproductive investment of the sexually dimorphic North Sea plaice (*Pleuronectes platessa* L.). J. Sea Res. 64, 85–92.
Vincenzi, S., Crivelli, A.J., Jesensek, D., De Leo, G.A., 2008. Total population density during the first year of life as a major determinant of lifetime body-length trajectory in marble trout. Ecol. Freshw. Fish 17, 515–519.
Watanabe, Y., Dings, G., Tian, Y., Tanaka, I., Stenseth, N., 2008. Determinants of mean length at age of spring spawning herring off the coast of Hokkaido, Japan. Mar. Ecol. Prog. Ser. 366, 209–217.
Wood, S., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. Ser. B 73 (1), 3–36.