



*J. Plankton Res.* (2016) 38(3): 589–603. First published online March 10, 2016 doi:10.1093/plankt/fbw020

# Habitat modelling of key copepod species in the Northwest Atlantic Ocean based on the Atlantic Zone Monitoring Program

SÉVERINE ALBOUY-BOYER<sup>1\*</sup>, STÉPHANE PLOURDE<sup>1</sup>, PIERRE PEPIN<sup>2</sup>, CATHERINE L. JOHNSON<sup>3</sup>, CAROLINE LEHOUX<sup>1</sup>, PETER S. GALBRAITH<sup>1</sup>, DAVE HEBERT<sup>3</sup>, GORDANA LAZIN<sup>3</sup> AND CAROLINE LAFLEUR<sup>1</sup>

<sup>1</sup>INSTITUT MAURICE-LAMONTAGNE, FISHERIES AND OCEANS CANADA, 850 ROUTE DE LA MER, CP 1000, MONT-JOLI, QC, CANADA G5H 3Z4,

<sup>2</sup>NORTHWEST ATLANTIC FISHERIES CENTRE, FISHERIES AND OCEANS CANADA, 80 EAST WHITE HILLS ROAD, PO BOX 5667, ST. JOHN'S, NL, CANADA A1C 5X1 AND

<sup>3</sup>BEDFORD INSTITUTE OF OCEANOGRAPHY, FISHERIES AND OCEANS CANADA, 1 CHALLENGER DRIVE, PO BOX 1006, DARTMOUTH, NS, CANADA B2Y 4A2

\*CORRESPONDING AUTHOR: boyerseverine6@gmail.com

Received October 23, 2015; accepted February 3, 2016

Corresponding editor: Roger Harris

Fourteen years of data collected by the Atlantic Zone Monitoring Program served to model the habitat of four dominant copepod species (*Calanus glacialis*, *Calanus hyperboreus*, *Calanus finmarchicus*, *Paracalanus* sp.) on the continental shelf and slope waters in the Northwest Atlantic Ocean. Generalized additive mixed models (GAMMs) were applied to abundance and presence–absence data for *C. hyperboreus*, *C. glacialis* and *Paracalanus* sp. and abundance for *C. finmarchicus*, to describe the optimal environmental productivity envelopes associated with the occurrence and/or the net productivity of these species. The models for *Calanus* species considered two main phases of their life cycle: (i) an active population growth phase dominated by early stages that occur primarily in surface layers, and (ii) a dormant phase dominated by overwintering stages generally found in deeper layers. GAMMs identified a marked contrast in environmental envelopes occupied by arctic and temperate species. Our analyses underline the importance of using data representative of all the copepodid developmental stages and occupied habitats in order to accurately model the distribution of *Calanus* species. The value of our models as tools to understand past events in the Northwest Atlantic or to predict future distributions of the species is also discussed.

**KEYWORDS:** *Calanus*; *Paracalanus*; generalized additive mixed models (GAMMs); habitat; CPR

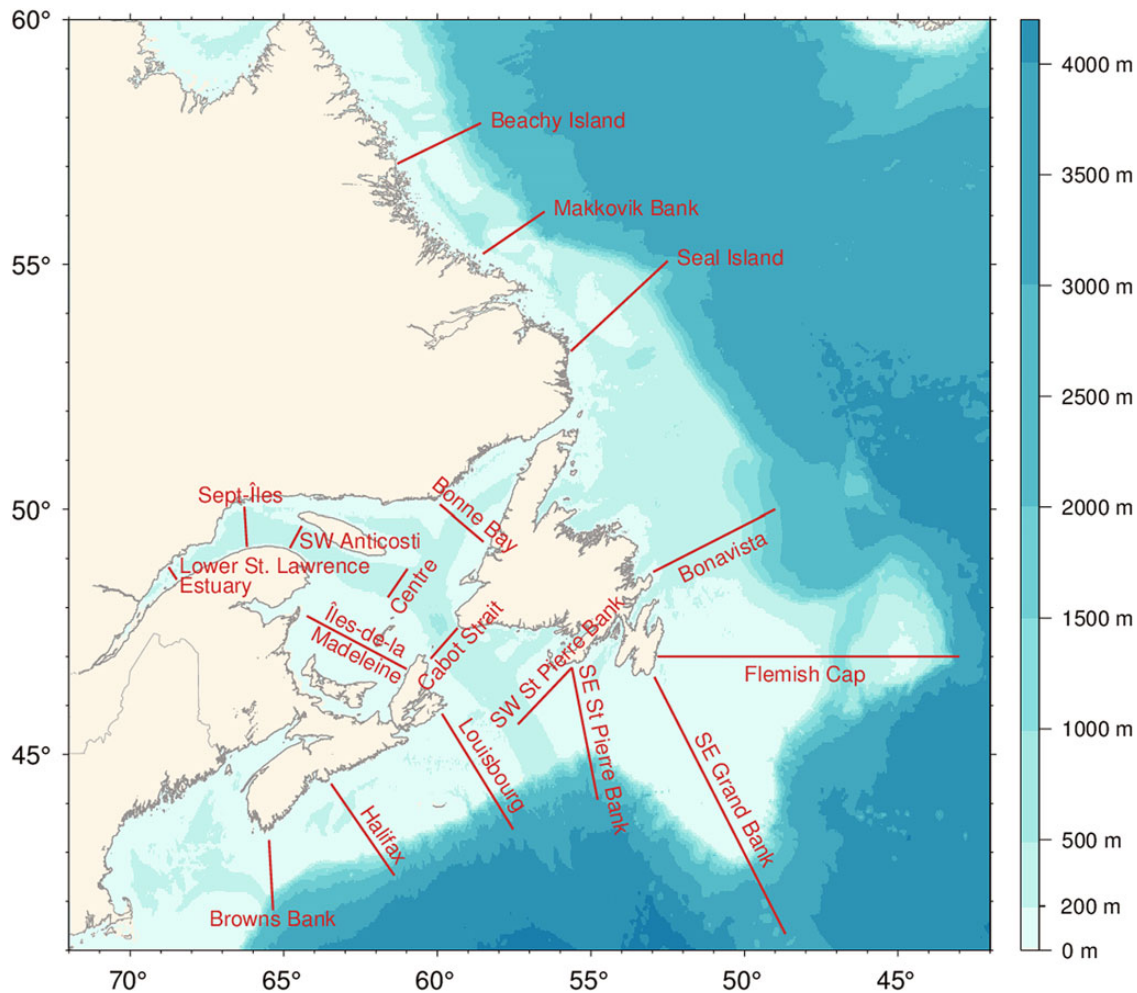
## INTRODUCTION

Modelling the processes that drive zooplankton species distribution is challenging because it results from both internal (organism level) and external controls (environment) (Planque *et al.*, 2011). To understand how environment (external controls) governs spatial and/or temporal zooplankton species distribution, several modelling approaches have been developed (for a complete review, see Chust *et al.*, 2014). In the North Atlantic ecosystem, the continuous plankton recorder (CPR) zooplankton time series provide the basis for development of models evaluating the role of environmental conditions on the distribution and occurrence of key copepod species and assessing the potential impacts of global warming (e.g. Helaouët and Beaugrand, 2007; Helaouët *et al.*, 2011, 2013, Chust *et al.*, 2014; Hinder *et al.*, 2014). These studies focus on the copepod *Calanus finmarchicus*, a keystone North Atlantic zooplankton species. They are based on the ecological niche concept [*sensu* Hutchinson (Hutchinson, 1957)] and provide comprehensive knowledge of the factors controlling its probability of occurrence at large spatial and temporal scales (Helaouët and Beaugrand, 2007; Beaugrand and Helaouët, 2008; Helaouët *et al.*, 2013).

The Atlantic Zone Monitoring Program (AZMP), conducted by the Department of Fisheries and Ocean (DFO) Canada since 1999, collects physical, chemical and biological data on phytoplankton and zooplankton communities over the whole water column (max depth: 1000 m for zooplankton samples) on the Newfoundland–Labrador Shelf (NLS), in the Gulf of St Lawrence (GSL) and on the Scotian Shelf (ScS) using standard protocols (Therriault *et al.*, 1998; Mitchell *et al.*, 2002). The program monitors environmental variations on seasonal and interannual scales over large spatial scales along oceanographic sections that cover regions ranging from the Labrador Shelf in the north to the lower St Lawrence estuary (LSLE) in the west and the southwestern ScS in the south (from 43° to 69°W; 41° to 58°N) (Fig. 1). Because of large-scale biogeographic conditions in this area, the zooplankton community is characterized by the presence of arctic (e.g. *Calanus hyperboreus*, *Calanus glacialis*), sub-arctic (e.g. *C. finmarchicus*) and temperate (e.g. *Paracalanus* sp.) species influenced by a broad range of water masses. The main components of the mesozooplankton community in these regions are from the genus *Calanus* which accounts for up to 70% of the total mesozooplankton biomass (Head *et al.*, 2003). *Calanus* species life cycles are characterized by a dormancy phase where copepodid stages (CIV–CVI depending on the species) overwinter at depths >100 m. Moreover, *C. glacialis* and *C. hyperboreus* have multi-annual life cycles (Kosobokova, 1999; Plourde *et al.*, 2003). Among

*Calanus* species, *C. finmarchicus* has been the subject of numerous studies on various aspects of its life history and demography in North Atlantic Ocean (see Melle *et al.*, 2014 and references therein). In the Northwest Atlantic Ocean, this calanoid copepod is a dominant component of the mesozooplankton biomass from winter until late summer in the Labrador Sea, on the continental shelf off Labrador and Newfoundland, in the GSL and ScS and in the Gulf of Maine/Georges Bank region (Meise and O'Reilly, 1996; Head *et al.*, 1999, 2003, 2013; Pepin *et al.*, 2005). The arctic *C. hyperboreus* is a key plankton species in the North Atlantic and Arctic Oceans (Hirche, 2013, and references therein) and is an important contributor to the zooplankton biomass in deep regions of the ScS, NLS and GSL (Sameoto and Herman, 1992; Head *et al.*, 2003; Plourde *et al.*, 2003). *Calanus glacialis* is endemic to Arctic seas (Plourde *et al.*, 2005; Darnis and Fortier, 2014) and is associated with shelf waters on the NLS (Head *et al.*, 2003; Pepin *et al.*, 2011) and ScS (Herman *et al.*, 1991; Sameoto and Herman, 1992) as far south as southwest Nova Scotia (Runge *et al.*, 1985). During the 1990s, increases in the abundance of the two arctic species *C. hyperboreus* and *C. glacialis* in the NLS and ScS areas were observed in CPR records and may have been associated with a rise in arctic water influx (Head and Sameoto, 2007; Head and Pepin, 2010). In parallel, an increase in the abundance of warm-water species at higher latitudes, such as *Paracalanus* sp., has also been described (Plourde *et al.*, 2014). *Paracalanus* is a ubiquitous genus found in most oceans (Razouls *et al.*, 2015). In the Gulf of Maine and on the ScS, this small copepod genus (identified as *Paracalanus parvus*) is considered among the dominant copepods by abundance in these regions (Johnson *et al.*, 2011), but there have been no comprehensive analyses of the habitat for this genus.

The AZMP database, in all its facets (sampling, spatial and temporal scales), represents a remarkable opportunity to describe the realized habitat of key copepod species over 17° of latitude and understand how environmental conditions, as bottom-up processes, could influence their productivity and distribution. The objective of our study was to describe the environmental factors controlling the productivity (abundance) and spatial distribution (occurrence) of *C. hyperboreus*, *C. glacialis*, *Paracalanus* sp. and the productivity of *C. finmarchicus* in the Northwest Atlantic Ocean using habitat models constructed with data representative of all the copepodid developmental stages and their habitat. These species were selected based on their distribution range and overall importance in the ecosystem. The AZMP region includes the northern limit of *Paracalanus* sp. distribution (Johnson *et al.*, 2011) and the southern edges for *C. hyperboreus* and *C. glacialis* (Head and Pepin, 2010), an important factor in habitat modelling based on occurrence data. *Calanus finmarchicus* was also



**Fig. 1.** Sampling areas of the AZMP.

included because of its ecological importance across the North Atlantic and because of its omnipresence in the AZMP region, its habitat was modelled only with abundance data. To determine the realized habitat of these four species, we developed a set of generalized additive mixed models (GAMMs) taking into consideration (i) the predominantly active and dormant development stages of *Calanus* species, (ii) pertinent physical and biological environmental variables to describe the habitat and (iii) temporal (seasonal and interannual) trends. GAMMs were performed to describe the environmental envelopes (habitat characteristics) constraining species productivity (abundance) and occurrence (presence/absence) in the AZMP region. Results are discussed in the context of previous studies describing the habitat of *Calanus* species and other copepod species in the North Atlantic Ocean and the potential of using our models to predict the effect of future climate changes on their distribution and productivity in the Northwest Atlantic Ocean.

## METHOD

### Sampling

Samples were collected as part of AZMP surveys from 1999 to 2012 in three major marine regions: the NLS, the GSL and the ScS. Physical, chemical and biological data were collected on 19 sections: 4 on the ScS, 8 on the NLS and 7 in the GSL (Fig. 1). A total of 4842 samples were available for the period 1999–2012, 917 for the ScS, 2019 for NLS and 1906 in the GSL. As mentioned in Pepin *et al.* (Pepin *et al.*, 2015), the three regions were not sampled simultaneously every year. NLS stations were sampled before or during the spring bloom (April–May), in November–December for the autumn surveys and two sections (Bonavista, Flemish Cap) were also sampled in summer (July, August), whereas Beachy Island, Makkovik Bank and Seal Island were surveyed only in summer (Fig. 1). In the ScS regions, surveys occurred in April–May, just after the bloom, and in

September–October for the autumn sampling. In contrast, in the GSL, spring cruises occurred in June, well after the spring bloom and in October–November for the autumn monitoring (for details, see Fig. 2 in Pepin *et al.*, 2015). Hydrographic profiles and water sampling were collected with a CTD-Rosette system following the AZMP protocol (for more details on sampling protocol and laboratory analyses, see Mitchell *et al.*, 2002). Zooplankton was sampled using 0.75 m diameter ring net with 202  $\mu\text{m}$  mesh nets towed vertically from 5 m above the bottom or to a maximum 1000 m to surface. The mesh size used and the targeted maximum sampling depth imply that the majority of the different copepodid stages (ca. active and dormant) of the *Calanus* species were sampled, the exception being for *Paracalanus* sp. CI–CIII because of its smaller body size (Melle *et al.*, 2014). Samples were preserved in 2% formaldehyde on deck. Species and stages were identified using sub-samples containing at least 200–400 organisms. Copepod abundance ( $\text{ind. m}^{-2}$ ) was determined and identification to copepodite stages was available for each species except *Paracalanus* sp. for which the combined abundance of stages CI–CVI was considered for analysis.

One objective of our study was to distinguish the active surface dwelling and deep-dwelling overwintering components of *Calanus* species populations. In general, the bulk of the actively growing component is associated with the food-rich upper layers (ca. 0–100 m), while the overwintering component seeks refuge in the deeper layers to minimize mortality during periods of adverse environmental conditions (Conover, 1988). The overwintering component of the arctic *C. glacialis* and *C. hyperboreus* is generally composed of stages CIV, CV and CVI (Conover, 1988; Plourde *et al.*, 2014), and the presence of these stages at the surface is restricted to the winter–spring period in the region (Plourde *et al.*, 2003; Head and Pepin, 2010). These characteristics mean that the combined CI–CIII abundance would represent a reliable estimate of populations developing in the surface layer, mainly during spring (March–May), while the combined CIV–V–VI abundance sampled in summer and autumn would be mostly associated with the overwintering component (Supplementary data, Fig. S1a and b). The situation is different for the sub-arctic *C. finmarchicus*. This species shows a predominant 1-year life cycle in the region with CV of the first generation being the dominant overwintering stage (Conover, 1988; Plourde *et al.*, 2001; Head and Pepin, 2007). This main pattern is accompanied by the potential for the production of secondary generations in late summer and autumn in response to environmental conditions (Head and Pepin, 2010; Maps *et al.*, 2010). Our AZMP data set generally conforms to this pattern (Plourde *et al.*, 2009; Pepin *et al.*, 2015): (i) the stage structure in spring (April–May) on the NLS is generally dominated by CVI with the remnant of

the CV overwintering stock and with few early stages, (ii) all developmental stages are abundant during spring (April) and summer (June, July) on the ScS, GSL and NLS, respectively, and (iii) the stage structure is dominated by CV in all regions in the autumn with the presence of early stages and CVI in relatively low abundance indicative of active production (Supplementary data, Fig. S1c). Therefore, we choose to group Stages CI–IV to represent the actively growing component, and Stages CV–CVI that would mainly represent the overwintering component of the population with the contribution of some active CV–CVI. This CV–CVI group also corresponds to stages identified to the species level in the CPR data set and used in previous *Calanus* species habitat modelling (e.g. Head and Pepin, 2010; Chust *et al.*, 2014).

As *Calanus* species have a complex life cycle with an active phase mainly in the upper 100 m and a dormancy phase at depths > 100 m, we used a variety of environmental indices from the AZMP database for a multivariate evaluation of environmental effects on zooplankton community structure (Pepin *et al.*, 2015). The pelagic habitat during the active phase was described with the following variables: average temperature and salinity in the 0–50 m depth layer (T0\_50, S0\_50), the integrated chlorophyll concentration 0–100 m (CHL0\_100) and the stratification index (STRAT) [difference in density between 50 and 5 m, a measure of the near-surface density gradient used in AZMP state of the ocean reporting and an indices previously used in AZMP framework (e.g. Pepin *et al.*, 2015 and references therein)]. In contrast, near bottom temperature (T\_NB) was considered as a descriptor of the environment during dormancy. Station depth (DEPTH) was also included as a general descriptor of the habitat (depth range) available to these species. With the aim to compare with CPR data, average temperature in the 0–5 m layer (T\_SURF) was also determined. We did not include latitude and longitude because T0\_50, S0\_50, T\_NB and STRAT had a strong latitudinal and longitudinal pattern in the studied area.

## Habitat modelling

Environmental variables considered in the habitat modelling were chosen because of known or anticipated effects on zooplankton (Pepin *et al.*, 2015). The correlations among all environment variables were evaluated using Pearson's correlation coefficients; variables with correlations exceeding 0.5 were not included simultaneously in models to control for collinearity (Wintle *et al.*, 2005; Heinänen *et al.*, 2008). For instance, T0\_50 and T\_SURF were highly correlated ( $r = 0.83$ ), but T0\_50 was considered to be a better index of the habitat occupied by actively growing *Calanus* species across their distribution range



(Gislason *et al.*, 2000; Plourde *et al.*, 2001, 2003; Basedow *et al.*, 2010). As a result, the explanatory variables kept for the habitat modelling were T0\_50, T\_NB, DEPTH, S0\_50, STRAT and CHL0\_100. Copepod abundances, STRAT, DEPTH and CHL0\_100 were ln-transformed (natural logarithm)  $[\ln(X + 1)]$  to minimize the effect of skewed data distributions.

A major issue concerning count data is that species may be rare or have a non-random distributions resulting in a higher proportion of zero values than expected according to a Poisson's distribution (Barry and Welsh, 2002). To address this problem, a two-step conditional approach was applied to the three species (*C. glacialis*, *C. hyperboreus* and *Paracalanus* sp.) with a high proportion of stations where they were absent (e.g. Borchers *et al.*, 1997; Heinänen *et al.*, 2008). In the first step, termed “abundance model”, absences (zero values) were excluded from the model. Abundance data were modelled assuming a Gaussian distribution (Barry and Welsh, 2002). In the second step, termed “occurrence model”, we modelled species occurrence by computing a presence/absence matrix (1, 0) following a binomial distribution. As *C. finmarchicus* was present in 97% of samples, its habitat description was modelled solely with abundance data, including the small proportion of stations where it was absent (abundance = 0).

### Generalized additive mixed models

To minimize the effect of the potential spatial autocorrelation along transect data, three depth classes were created based on station bottom depth: <200, 200–500 and

>500 m. All environmental and zooplankton variables were averaged for each of these bathymetric categories on each section for each mission, resulting in 240, 450 and 316 data for the ScS, NLS and GSL, respectively. To take into account temporal patterns and especially the strong seasonal variability between years, the habitat modelling was performed with GAMMs including year and season (spring, summer and autumn) as nested random effects. Like GAMs, GAMM includes non-parametric terms and is data-driven rather than being restricted to an *a priori* defined error distribution (Yee and Mitchell, 1991). It allows the capture of non-linear asymmetric or bimodal biological responses and unlike GAMs, the mixed model formulations allow treatment of autocorrelation and repeated measure situations (Wood, 2006). The choice to include a mixed effect (=random effect) in our models was a way to avoid the potential influence of a long-term trend in environmental conditions that could induce bias for predictions made in the past or in the future. To avoid over-fitting and complex multimodal response curves, the estimated degrees of freedom ( $k$ ) for the predictors were restricted from 1 to 5 and the smoothing functions for the different predictors were based on the thin plate regression spline because it is computationally stable, avoiding the problem of knot placement and suitable for any number of predictor variables (Wood, 2006). GAMMs were performed with the “*gam4*” package in R (R Core Team, 2014) using restricted maximum likelihood (REML) for smoothness selection (Wood and Scheipl, 2014). The GAMM produces fitted values of the response variable (here abundance or occurrence probability) using an identity or a logit link function, respectively.

Table I: GAMMs results for *C. glacialis*

CI–CIII	Abundance model ( $n = 415$ ; $R^2 = 0.38$ )			Occurrence model ( $n = 868$ ; TSS = $0.52 \pm 0.06$ )		
	Significance	Smooth order	Partial res. >0	Significance	Smooth order	Partial res. >0
T0_50	***	3.37	<4°C	***	1	<5°C
T_NB	***	1	<3°C	***	1	<4°C
S0_50	***	2.68	>31	***	3.53	31.8–34.5
IDEPH	–	–	–	–	–	–
ISTRAT	–	–	–	–	–	–
ICHL0_100	***	2.86	>4.8	*	1	>4
CIV–CVI	Abundance model ( $n = 698$ ; $R^2 = 0.31$ )			Occurrence model ( $n = 870$ ; TSS = $0.48 \pm 0.07$ )		
	Significance	Smooth order	Partial Res. > 0	Significance	Smooth order	Partial res. > 0
T0_50	***	1	<5°C	***	1	<5°C
T_NB	**	1.60	<3°C	***	3.07	3.5–9°C
S0_50	–	–	–	***	2.47	>33
IDEPH	***	3.70	5.3	–	–	–
ISTRAT	**	3.35	0.7–1.5	***	1	>0.7
ICHL0_100	***	1	<4.3	***	1	<4

$R^2$ , true statistic skill (TSS) and number of data used to model habitat ( $n$ ) are noted. For each environmental variable, its significance, smooth order and influence (Res. value) on the abundance or occurrence probability are summarized.

\*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , ns, non-significant; –, not selected.

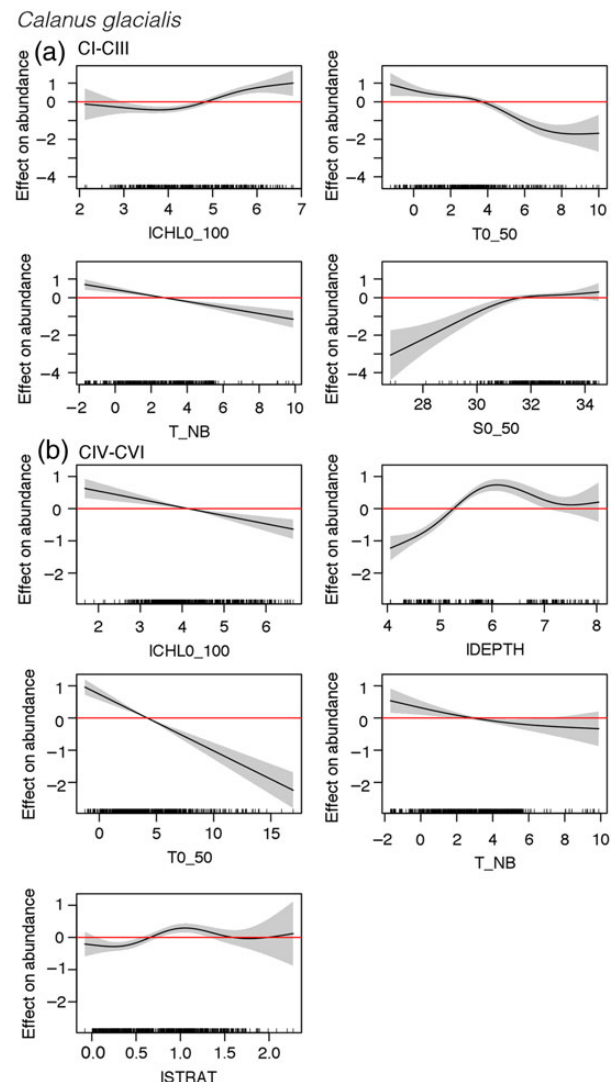
After data transformation (removing 0 and NA/transform abundance in binomial matrix), 80% of the data set was randomly selected for the model calibration and the remaining 20% was used for model validation. For the occurrence models, species prevalence (per cent of presence and absence in the matrix) was taken into account in the sub-sampling. For each copepod indices models, an automated selection of environmental variables was performed from a global model including the six explanatory variables chosen by scatterplot explorations (T0\_50, T\_NB, S0\_50 and log-transformed variables: DEPTH, STRAT, CHL0\_100). GAMM selection was undertaken by comparing all computed GAMMs including all possible combination of explanatory variables. The selection of the best models was based on the lowest AIC (Akaike information criterion) values and only the models within 2 AIC ( $\Delta < 2$ ) units of the most parsimonious model (i.e. the model with the lowest AIC) were retained (Grueber *et al.*, 2011). This automated selection model was performed with the “dredge” and “uGamm” functions in MuMIn package (Bartón, 2014). The best model chosen was the one with the best predictive power during the validation process (the 20% of data). The best predictive model for the abundance model was the one with the best Spearman’s correlation between predicted (model) and observed data (20% of data set). For occurrence models, the efficiency of the models was estimated by the True Skill Statistic ( $TSS = \text{sensitivity} + \text{specificity} - 1$ ) (Allouche *et al.*, 2006) using the package “PresenceAbsence” (Freeman and Moisen, 2008). This cross-validation procedure was repeated 99 times in order to avoid bias in the random sampling. For the 99 iterations, we have estimated the percentage of selection for each variable and the average number of variable selected. The final selected model was the one with the average number of the selected variable over time for the 99 iterations and was rerun with all the data. For occurrence models, a mean TSS was calculated on 100 runs for each selected model by species. The statistical framework is detailed in Supplementary data, Fig. S2.

## RESULTS

In GAMMs, partial response curves can be interpreted by examining residual values. If residual values are  $>0$ , including confidence interval (grey area), then that predictor has a positive influence on the response variable (here abundance or occurrence probability) at that particular predictor value; the curve shows the effect of the individual predictor (e.g. T0\_50) on the model taking into account the effects of the other variables in the model (Papastamatiou *et al.*, 2013). Although analyses were performed with some ln-transformed explanatory

variables (STRAT, CHL0\_100 and DEPTH = ISTRAT, ICHL0\_100 and IDEPTH), values were back-transformed in the text to better describe habitat characteristics.

The abundance of CI–CIII *C. glacialis* was governed by four environmental variables, (i.e. T0\_50, T\_NB, S0\_50 and ICHL0\_100) with 38% of the variance explained by the model ( $R^2 = 0.38$ , Table I). Greater abundance was associated with T0\_50 and T\_NB lower



**Fig. 2.** GAMMs showing the influence of predictor variables on *C. glacialis* abundance (log-transformed). (a) CI–CIII model, and (b) CIV–CVI model. The ranges of environmental variables are represented on the x-axis and the influence of the predictor holding constant the other variables in the model on the y-axis. Zero on the y-axis (red line) means the environmental variable has no effect. The degree of smoothing is indicated in Table I. Shaded areas represent 95% confidence interval for the response curves. Ticks on the x-axis indicated the data observations. (Colour version of the figure is available online.)

than 4°C and 3°C, respectively; S0\_50 higher than 31 and CHL0\_100 above 120 mg m<sup>-2</sup> (ICHL0\_100 > 4.8) (Fig. 2a). The same variables were selected in the occurrence model with an overall good accuracy (TSS = 0.52 ± 0.06), although ICHL0\_100 was less influential and the positive effect of S0\_50 was restricted to a smaller range (31.8–34.5) (Supplementary data, Fig. S3a; Table I). In the CIV–CVI *C. glacialis* abundance model, five variables were selected, accounting for 31% of the variance ( $R^2 = 0.31$ , Table I). Unlike the CI–CIII model, S0\_50 was not selected, but ISTRAT and IDEPTH were significant. Greater abundance of *C. glacialis* CIV–CVI was associated with temperatures lower than 5°C (T0\_50 and T\_NB), CHL0\_100 lower than 70 mg m<sup>-2</sup> (ICHL0\_100 < 4.3), depth > 200 m, (IDEPTH > 5.3) and a stratification index between 1 and 3.5 kg m<sup>-3</sup> (0.7 > ISTRAT > 1.5) (Fig. 2b). The occurrence model of CIV–CVI *C. glacialis* showed a higher probability of occurrence for similar conditions described above except for depth (not selected) and S0\_50 higher than 33 (Supplementary data, Fig. S3b). The accuracy of the model was good (TSS = 0.48 ± 0.07) (Table I).

Abundances of *C. hyperboreus* were well described by selected GAMMs with good explanatory power ( $R^2 = 0.45$  and 0.52 for CI–CIII and CIV–CVI, respectively) (Table II). Abundance of CI–CIII was greater at CHL0\_100 higher than 150 mg m<sup>-2</sup> (ICHL0\_100 > 5), S0\_50 between 30 and 32.5, T0\_50 and T\_NB below 5°C and 2°C, respectively, and stations deeper than 300 m (IDEPTH > 5.8). Higher abundance was also associated with T\_NB > 7°C, although this pattern should be considered as marginal because it was based on few observations (Fig. 2a). The

occurrence model for CI–CIII showed the lowest accuracy for all models, with a mean TSS of 0.34, and did not include T0\_50 (Table II). However, the effects of selected environmental variables were similar to those observed in the abundance model. The abundance model of *C. hyperboreus* CIV–CVI considered all environmental variables as significant with the exception of T\_NB (Fig. 3b; Table II). Similar patterns to those observed for CI–CIII were apparent, but the stratification index was selected with a significant positive effect on abundance. Accuracy for the CIV–CVI occurrence model was good (TSS = 0.43 ± 0.08). The occurrence probability of CIV–CVI *C. hyperboreus* was higher for T\_NB above 3.5°C, S0\_50 lower than 32 and depth > 300 m (IDEPTH = 5.7). No significant effect of ICHL0\_100 was detected. Stratification index and T0\_50 were not selected in the two occurrence models for this species (Supplementary data, Fig. S4).

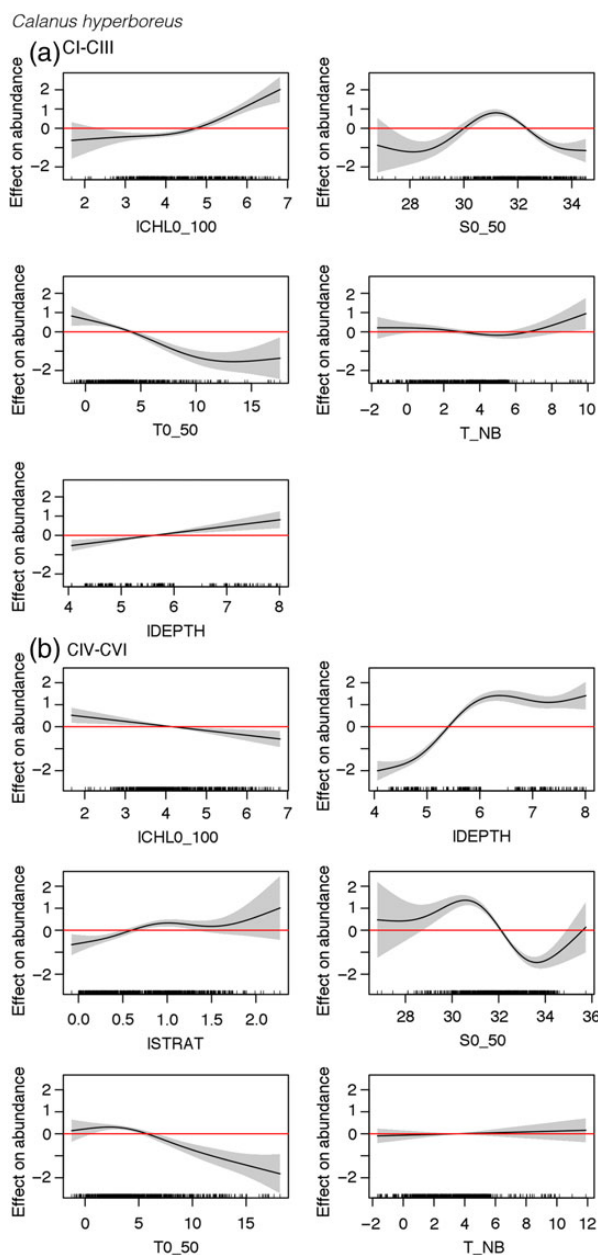
GAMMs for CI–CIV and CV–CVI *C. finmarchicus* abundance resulted in 15 and 25% of total variance explained, respectively (Table III). In CI–CIV GAMM, all environmental variables were included and were significant with the exception of ICHL0\_100. T0\_50 showed a positive effect on *C. finmarchicus* CI–CIV abundance between 3 and 11°C, S0\_50 between 29 and 33, depths > 300 m (IDEPTH = 5.8) and for a stratification index lower than 1 kg m<sup>-3</sup> (ISTRAT < 0.7) (Fig. 4a). In contrast, all environmental indices were selected and highly significant for the CV–CVI abundance model (Fig. 4b). Greater CV–CVI abundance was associated with a similar depth and T0\_50 ranges to that of the younger stages with S0\_50 above 32, CHL0\_100 above 54 mg m<sup>-2</sup> and stratification index lower than 1 kg m<sup>-3</sup> (Table III).

Table II: GAMMs results for *C. hyperboreus*

	Abundance model ( $n = 572$ ; $R^2 = 0.45$ )			Occurrence model ( $n = 870$ ; TSS = 0.34 ± 0.06)		
	Significance	Smooth order	Partial res. > 0	Significance	Smooth order	Partial res. > 0
CI–CIII						
T0_50	***	2.90	<5°C	–	–	–
T_NB	*	2.87	<2 to >7°C	***	2.74	<6°C
S0_50	***	3.85	30–32.5	***	2.76	30–33
IDEPTH	***	1	>5.8	***	3.64	5.2–7
ISTRAT	–	–	–	–	–	–
ICHL0_100	***	2.70	>5	*	2.14	>4.5
CIV–CVI						
	Abundance model ( $n = 786$ ; $R^2 = 0.52$ )			Occurrence model ( $n = 870$ ; TSS = 0.43 ± 0.08)		
	Significance	Smooth order	Partial res. > 0	Significance	Smooth order	Partial res. > 0
T0_50	***	2.96	<5.5°C	–	–	–
T_NB	ns	1	No relationship	***	1	>3.5°C
S0_50	***	3.90	<32	**	1	<32
IDEPTH	***	3.86	>5.5	***	1	>5.7
ISTRAT	***	3.20	>0.7	–	–	–
ICHL0_100	**	1	<4	n.s.	1	No relationship

$R^2$ , true statistic skill (TSS) and number of data used to model habitat ( $n$ ) are noted. For each environmental variable, its significance, smooth order and influence (Res. value) on the abundance or occurrence probability are summarized.

\*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , n.s., non-significant; –, not selected).



**Fig. 3.** GAMMs showing the influence of predictor variables on *C. hyperboreus* abundance (log-transformed). (a) CI–CIII model, and (b) CIV–CVI model. The ranges of environmental variables are represented on the x-axis and the influence of the predictor holding constant the other variables in the model on the y-axis. Zero on the y-axis (red line) means no effect on the environmental variable. The degree of smoothing is indicated in Table II. Shade area represents 95% confidence interval for the response curves. Ticks on the x-axis indicated the data observations. (Colour version of the figure is available online.)

*Paracalanus* abundances were driven mainly by T0\_50 and S0\_50 and the abundance model was robust with 57% of the total variance explained ( $R^2 = 0.57$ ) (Table IV). *Paracalanus* abundance was positively influenced by T0\_50

higher than 7°C and S0\_50 m between 30 and 33 (Fig. 5). Abundance was also positively influenced, but to a lesser extent, by integrated chlorophyll (0–100 m)  $> 54 \text{ mg m}^{-2}$  (ICHL0\_100  $> 4$ ) and T\_NB above 4°C, while IDEPTH and ISTRAT were not selected (Table IV, Fig. 5). The accuracy of the occurrence model was very good (mean TSS =  $0.73 \pm 0.05$ ) even if only three of the five selected predictors were significant (T0\_50, S0\_50 and ISTRAT) (Table IV). In contrast to the abundance model, two non-significant variables (T\_NB and ICHLO\_100) were conserved in the occurrence model. The effect of T0\_50 was similar to the abundance model (Supplementary data, Fig. S5). In both models, the effect of ICHLO\_100 was at best weak (Table IV).

## DISCUSSION

### Modelling approach

The objective of the modelling framework developed in this study was to apply an original approach in habitat modelling and compare our results with existing studies. Beyond the robustness of the statistical approach developed in our study (see Supplementary data, Fig. S2), the need to determine how various stage-components of the *Calanus* species would respond to changing environmental conditions guided our analyses. Models developed for early stages served to characterize the habitat favouring cohort development and describe a more direct and immediate effect of the environment on the population productivity. The situation was more complex for the older stages. Even if this group could at times be active in the surface layer for short periods, we have considered that the response described in our models would be dominated by overwintering stages and reflect the resilience of *Calanus* species populations to adverse environmental conditions during overwintering. Given the arrested/slowed development during overwintering and the multi-year life cycle of *C. glacialis* and *C. hyperboreus*, our results should describe how overwintering populations coped with variations in habitat characteristics occurring over a much larger spatial and temporal scale than the response of early stages to a direct and immediate effect on species. We recognize that the stage groups used to distinguish the active and overwintering components of *Calanus* species populations were based on general characteristics of their life cycle and would be unlikely to account for variations observed either across the region or over the years considered in our study. This is particularly true for *C. finmarchicus*, a species that demonstrates a more plastic phenology capable of producing more than one generation under warmer environmental conditions, which could result in a longer active period for at least a fraction of its population (Plourde et al., 2009; Pepin et al., 2015).



Table III: GAMMs results for *C. finmarchicus*

	CI-CIV ( $n = 870$ ; $R^2 = 0.15$ )				CV-CVI ( $n = 870$ ; $R^2 = 0.25$ )		
	Significance	Smooth order	Res. value > 0		Significance	Smooth order	Res. value > 0
T0_50	***	3.83	3–11°C	***	3.56	5–11°C	
T_NB	***	3.76	0–3°C; 6–10°C	***	3.01	>4°C	
S0_50	***	3.42	29–33	***	1	<32	
IDEPH	***	1	>5.8	***	3.58	5.3–7.8	
ISTRAT	***	1	<0.7	***	1	<0.7	
ICHL0_100	n.s.	1	No relationship	***	1	<4	

$R^2$ , and number of data used to model habitat ( $n$ ) are noted.

For each environmental variable, its significance, smooth order and influence (Res. value) on the abundance are summarized.

\*\*\* $P < 0.001$ , n.s., non-significant.

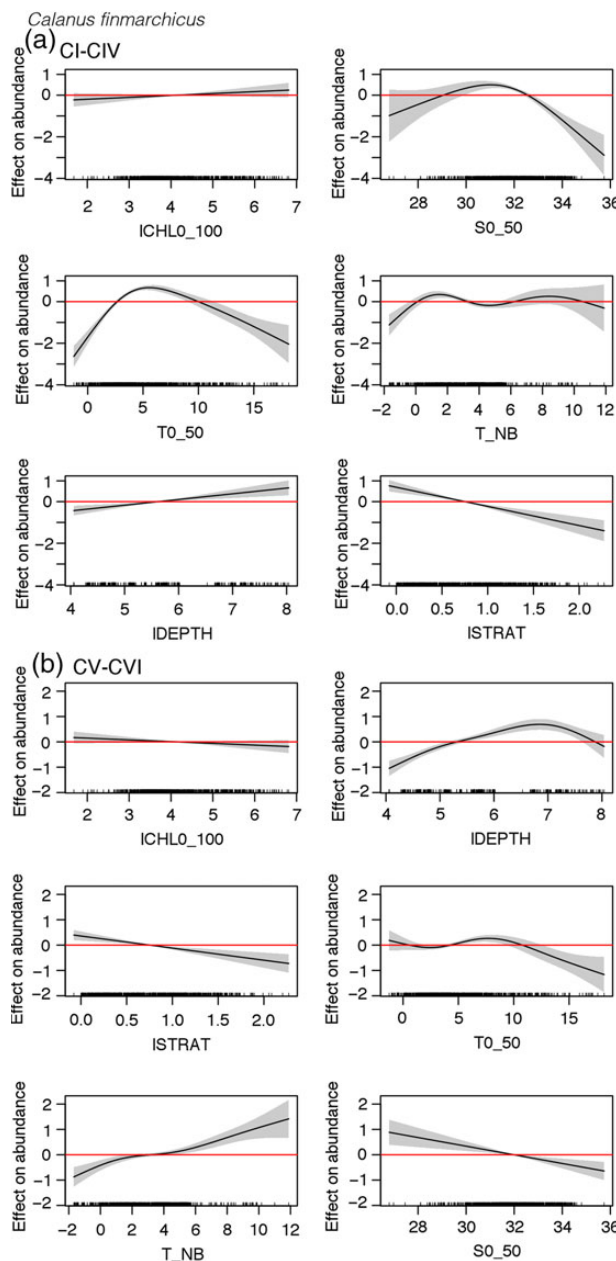
Our depth-integrated abundance data allowed the development of models describing the environmental conditions associated with high population abundance, despite the fact that our data did not encompass the entire distributional range of our target species (northern edge for *Paracalanus* sp. and southern limits for the two arctic species *C. glacialis* and *C. hyperboreus*). Interestingly, environmental variables selected to predict abundance were not always identical to those selected to determine the occurrence probability of the species. The best example was for *C. hyperboreus* models in which temperature (0–50 m) was not selected with the occurrence data, whereas it was highly significant in abundance models (Table II). The difference in variables explaining species' occurrence and abundance indicates that results for these models can be interpreted separately (Cunningham and Lindenmayer, 2005). In general, occurrence models would predict an extreme response of the species population to environmental conditions resulting in their absence (e.g. extinction), while abundance models predict the productivity response of species. This would explain why environmental windows could be wider for occurrence models (e.g. a difference of 2°C for T0\_50 range for *Paracalanus* sp. model). However, according to Barry and Welsh (Barry and Welsh, 2002), the two models can also be combined, allowing the prediction of the abundance of the species “conditioned” by its probability of occurrence and therefore, representing an interesting tool to predict future distribution and productivity of species.

In general, the different models selected provided reliable descriptors of the pattern of occurrence and abundance that can be considered as robust. Occurrence models showed relatively good accuracy, except for *C. hyperboreus* (TSS from 0.34 to 0.43) (Allouche et al., 2006). Lower accuracy for this species can be explained by the choice of a threshold that maximized the sum of sensitivity and specificity (Freeman and Moisen, 2008). Sensitivity is the proportion of observed presence correctly predicted (true positive) and specificity is the proportion of observed absence correctly predicted (true negative) (Allouche et al., 2006),

meaning that the average of the error rate for positive observations and the error rate for negative observations were minimized and tend to overestimate the prevalence of rare species while underestimating the prevalence of common species. Accordingly, the best TSS was observed for models applied to rarer species such as *Paracalanus* sp. (TSS = 0.73) and *C. glacialis* (TSS around 0.5). Except for *C. finmarchicus* models, explained variances of abundance models were relatively high ( $R^2$  from 0.31 to 0.57). The lower explained variance ( $R^2 = 0.15$  for the CI-CIV model, 0.25 for the CV-CVI model) observed for *C. finmarchicus* abundance models is probably partly caused by its ubiquity in the region (the species was recorded in 97% of data) and the fact that environmental conditions across the region during the periods considered were not extreme enough to hinder the ability of the species to be productive during its active phase or survive during the overwintering period.

### Habitat modelling versus physiological and ecological knowledge of target species

High abundance of the active component (CI-CIII) of the two arctic species, *C. glacialis* and *C. hyperboreus*, was associated with T0\_50 lower than 4–5°C and high phytoplankton biomass typical of the spring environmental conditions in the AZMP region. These results confirm the affinity of these arctic species to low temperatures when active in the surface layer (Kosobokova, 1998; Niehoff and Hirche, 2005; Chust et al., 2014) and the general dependency and synchrony of the development of young stages to the spring phytoplankton bloom (Hirche and Kattner, 1993; Plourde et al., 2005; Broms et al., 2009; Morata and Soreide, 2015). Very similar environmental variables with similar effects on the abundance or occurrence of the active stages were selected in both species. However, the selection of the location depth in models of CI-CIII *C. hyperboreus* corroborates the association of this species with deeper habitats (Head et al., 2003; Plourde et al., 2003), whereas *C. glacialis* is more associated with shelf/shallow



**Fig. 4.** GAMMs showing the influence of predictor variables on *C. finmarchicus* abundance (log-transformed). (a) CI–CIV model, and (b) CV–CVI model. The ranges of environmental variables are represented on the x-axis and the influence of the predictor holding constant the other variables in the model on the y-axis. Zero on the y-axis (red line) means no effect on the environmental variable. The degree of smoothing is indicated in Table III. Shade area represents 95% confidence interval for the response curves. Ticks on the x-axis indicated the data observations. (Colour version of the figure is available online.)

regions (Head *et al.*, 2003; Pepin *et al.*, 2011). The active component of the sub-arctic *C. finmarchicus* was more strongly associated with spring–summer transition conditions well after the spring bloom in the GSL (Plourde

*et al.*, 2001) and on NLS (Pepin *et al.*, 2013). *Paracalanus* sp. abundance and occurrence were positively influenced by warm, post-bloom autumn conditions found in the region. Surprisingly, while Davis (Davis, 1987) noted that *Paracalanus* sp. seems to be associated with stratified areas and likely found above the thermocline on Georges Bank, the stratification index and the depth were not selected to explain its productivity. *Paracalanus* sp. usually dominates zooplankton community in shelf regions during the autumn in the North Atlantic Ocean (Labat *et al.*, 2009; Head and Pepin, 2010; McGinty *et al.*, 2011) but also occurs in deeper waters located in offshore areas (Pepin *et al.*, 2015).

Overwintering habitats were somewhat similar for the three *Calanus* species with bathymetry (DEPTH) as a key characteristic, as high abundance and occurrence probabilities were associated with depths typical of deep channels and basins located on the continental shelf or in deeper slope waters. The overwintering habitat of *C. glacialis* was characterized by cold and shelf/shallow waters, whereas *C. hyperboreus* and *C. finmarchicus* overwintering stages were more associated with deeper habitats typical of the deep channels in the GSL and to slope waters. Unlike its two congeners, overwintering of *C. glacialis* appears favoured in relatively cold environmental conditions typical of the arctic shelves, as it is known to overwinter in regions where near bottom temperatures are as low as 0°C (Kosobokova, 1999). In our study, high occurrence probability of *C. glacialis* overwintering stages was associated with relatively high T\_NB (3.5–9°C), but its abundance was low at the upper end of this range, as suggested by the negative effect of T\_NB above 3°C in the abundance model. This pattern suggests that *C. glacialis* is likely maintained in warmer regions such as the GSL and on ScS through advection from cooler northern areas such as NLS where conditions are more suitable for this arctic expatriate. In contrast, the absence of a significant effect of T\_NB on *C. hyperboreus* CIV–CVI abundance suggests that this species would be more tolerant to variations in temperature in its overwintering habitat. *Calanus hyperboreus* shows physiological rate responses (e.g. respiration) to temperature that are more similar to those of *C. finmarchicus* rather than *C. glacialis*, and its larger body size likely enables this species to tolerate much warmer overwintering conditions with success (Maps *et al.*, 2014). This tolerance to warmer conditions could explain its wider spatial distribution and its presence in regions where *C. glacialis* is seldom observed, such as the Central Labrador Sea (Head *et al.*, 2003) or the GSL (Plourde *et al.*, 2003). The sub-arctic species *C. finmarchicus* is broadly present in the whole study area with its overwintering component occurring principally at locations with depths in excess of 200 m and temperatures generally ranging from 5 to 11°C (Plourde *et al.*, 2001; Head and

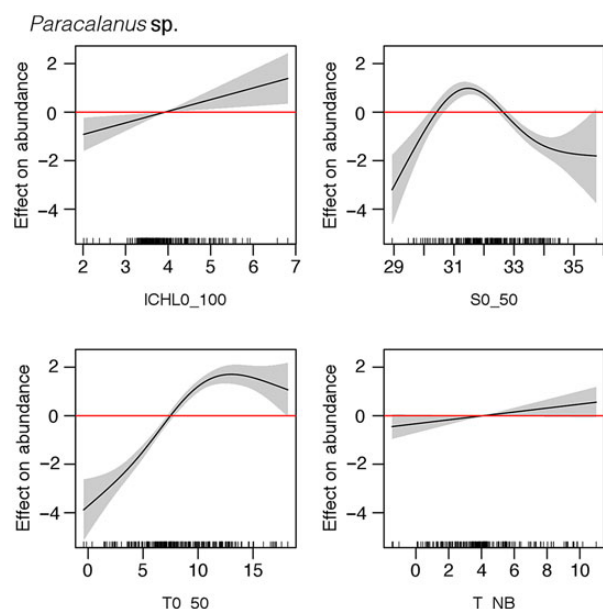
Table IV: GAMMs results for *Paracalanus sp*

	Abundance model ( $n = 236$ ; $R^2 = 0.57$ )			Occurrence model ( $n = 870$ ; $TSS = 0.73 \pm 0.05$ )		
	Significance	Smooth order	Partial res. > 0	Significance	Smooth order	Partial res. > 0
T0_50	***	3.35	>7°C	***	3.62	>5°C
T_NB	n.s.	1	>4°C	n.s.	1	No relationship
S0_50	***	3.64	30–33	***	3.19	Peak at 32
IDEPH	–	–	–	–	–	–
ISTRAT	–	–	–	***	3.25	<0.6
ICHL0_100	**	1	>4	n.s.	1	>4

$R^2$ , true statistic skill (TSS) and number of data used to model habitat ( $n$ ) are noted.

For each environmental variable, its significance, smooth order and influence (Res. value) on the abundance or occurrence probability are summarized.

\*\*\* $P < 0.001$ , \*\* $P < 0.01$ , ns, non-significant; –, not selected.



**Fig. 5.** GAMMs showing the influence of predictor variables on *Paracalanus sp.* abundance (log-transformed). The ranges of environmental variables are represented on the x-axis and the influence of the predictor holding constant the other variables in the model on the y-axis. Zero on the y-axis (red line) means no effect on the environmental variable. The degree of smoothing is indicated in Table IV. Shade area represents 95% confidence interval for the response curves. Ticks on the x-axis indicated the data observations. (Colour version of the figure is available online.)

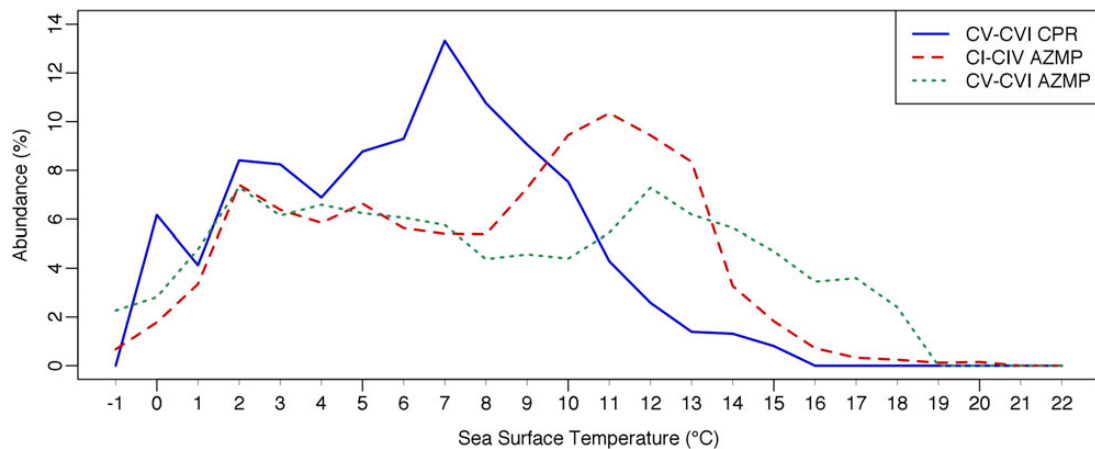
Pepin, 2007). According to these studies, our model showed that abundance of *C. finmarchicus* overwintering stages was positively influenced by T\_NB above 4°C.

Temperature was without doubt the most important driver allowing the discrimination of our four target species. Temperature ranges corresponding to relatively high species abundance and occurrence probabilities determined by our GAMMs were quite similar to those observed in others habitat studies (e.g. Carstensen et al., 2012; Helaouët et al., 2013). More importantly, temperatures detrimental for *Calanus* species abundance (negative effect) identified in our models were also similar to those negatively affecting various

individual physiological processes such as egg production (Pasternak et al., 2013), respiration/excretion (Alcaraz et al., 2014; Maps et al., 2014) and ingestion (Møller et al., 2012). For *C. finmarchicus*, our results also correspond to the general effect of elevated temperature on mortality and survival rates, a key parameter controlling population overall abundance and productivity in the AZMP region (Plourde et al., 2009). However, the direct effects of physical environmental conditions on the physiology and overall life history could not be the sole mechanisms controlling the seasonal succession of the three *Calanus* congeners and the delay of *C. finmarchicus* recruitment relative to the spring bloom (Broms and Melle, 2007; Broms et al., 2009). Intraguild competition for food and/or predation on *C. finmarchicus* early stages (eggs, nauplii) by its two arctic congeners arising early to the surface layer in spring could also contribute to the delay of recruitment observed for *C. finmarchicus* on the NLS and in the GSL (Head et al., 2000; Irigoien and Harris, 2006; Melle et al., 2014).

### *C. finmarchicus* habitat: AZMP versus CPR data

Most of the niche or habitat modelling of marine zooplankton in the North Atlantic Ocean has been performed using CPR and surface bio-physical environmental data covering a large horizontal spatial scale but sampling a small fraction of the zooplankton vertical habitat (e.g. Helaouët and Beaugrand, 2007; Beaugrand and Helaouët, 2008; Chust et al., 2014; Hinder et al., 2014). Our habitat models based on depth-integrated *C. finmarchicus* abundance data and environmental indices showed that its abundance is negatively affected by T0\_50 > 10–11°C (Table III, Fig. 4), an upper temperature threshold similar to the one quantified using CPR abundance and surface temperature data (see Fig. 16 in Helaouët and Beaugrand, 2007). However, comparing results obtained with data collected with such highly different sampling protocols could be misleading. The AZMP data set corresponds to the Longhurst's Northwest Atlantic shelves province (Longhurst, 1998), a region characterized



**Fig. 6.** Comparison of per cent relative average abundance of *C. finmarchicus* as a function of sea surface temperature 1°C bins obtained with the CPR and AZMP data sets. CV–VI CPR: data adapted from Fig. 5c in Helaouët and Beaugrand (Helaouët and Beaugrand, 2007). AZMP: CI–IV and CV–VI stages groups used in habitat models are presented separately. (Colour version of the figure is available online.)

by a large seasonal/spatial surface temperature gradient (−1.0–20°C) encompassed in the core distribution range of *C. finmarchicus* in the North Atlantic (Helaouët and Beaugrand, 2007). Based on these characteristics, we compared our results by plotting the percentage of CI–IV and CV–VI abundance observed against 1°C surface temperature bins (Fig. 6), a representation similar to the one made with CV–VI sampled by the CPR at the surface in Fig. 5c in Helaouët and Beaugrand (Helaouët and Beaugrand, 2007). Both data sets showed a similar detrimental effect of surface temperature lower than 2°C on abundance. However, our comparison revealed that depth-integrated CI–IV (surface dwellers) and CV–VI (also includes overwintering stages) abundances were negatively affected at a much higher surface temperature (14–15°C) than CPR surface abundance data (11–12°C) (Fig. 6), resulting in a more restricted optimal surface temperature envelope obtained with CPR surface abundance data. The most parsimonious reason explaining this difference between optimal temperature envelopes among these large data sets would be active avoidance of surface temperature > 12°C by the species, a behaviour that could affect CPR surface abundance but not depth-integrated abundance estimates. Temperature could affect the vertical distribution of the active *C. finmarchicus* as the species avoids highly stratified and warmer surface conditions and seeks refuge below the thermocline in cooler habitats (Williams, 1985; Basedow *et al.*, 2010; Jónasdóttir and Koski, 2011). Therefore, the sharp decrease in the occurrence probability of *C. finmarchicus* at surface temperatures higher than 12°C predicted by CPR-based models might be explained by this species avoiding sub-optimal surface temperature in addition to the negative impacts of warm conditions on individual and population processes. We

recognize that this simple comparison used data collected across different temporal and spatial scales, but it should be robust because of the relatively large number of observations and environmental range considered in the AZMP data set. Our comparison does not imply that reliable relative seasonal abundance patterns of *C. finmarchicus* active component cannot be derived from CPR data (i.e. Kane, 2005; Helaouët *et al.*, 2015), but rather points out that CPR surface data could overestimate the impacts of warmer surface conditions on *C. finmarchicus* abundance relative to depth-integrated sampling. We suggest that this issue warrants more attention and more quantitative comparison studies that should also consider other key zooplankton taxa.

As studies performed with the CPR database imply that the determination of ecological niches is based on a small fraction on the habitat used by this species, the effect of warming on *C. finmarchicus* occurrence might have been overestimated in a given geographical location, especially at the southern limit of its distribution. Indeed, there was a very high abundance of *C. finmarchicus* early stages in spring 2013 in the Gulf of Maine (southern edge of *C. finmarchicus* distribution) resulting from a high abundance of overwintering stages in the autumn of 2012, a year with record—high surface temperatures (i.e. 4–5°C warmer than the seasonal average) (Runge *et al.*, 2015). Application of our modelling approach at a regional scale could allow us to take into account the life cycle of the species, especially the active response of the species to avoid warm surface water, and seems to be complementary to CPR-data studies as (i) our models are made for active (CI–CIV) and overwintering stages (CV–CVI), and (ii) our modelling allows us to predict the productivity of the species.



## CONCLUSION

The habitat models developed for *Calanus* species in this study focused on two distinct phases of the life cycle: (i) the active population growth phase dominated by early stages in surface layers, and (ii) the dormant phase dominated by overwintering stages in deeper layers. The analysis highlights the importance of distinguishing between different developmental stages to accurately identify the influence of critical habitats features on the distribution of *Calanus* species. The statistical modelling approach used in our study provides a basis both for predicting future copepod distributional shifts from predictions of environmental change and for inferring past copepod distribution from environmental data sets in the absence of adequate past zooplankton data. Our models should provide useful foundational tools for the development of ecosystem-based advice about the role of environmental influence on past events [e.g. understanding the early 1990s collapse of groundfish stocks observed in the whole AZMP domain (Morissette et al., 2009; Dawe et al., 2012; Buren et al., 2014) as well as forecast the potential effects of climate change in Atlantic Canadian waters.

## SUPPLEMENTARY DATA

Supplementary data can be found online at <http://plankt.oxfordjournals.org>.

## FUNDING

The present study was funded by DFO's Aquatic Climate Change Adaptation Services Program (ACCASP) and a post-doctoral NSERC (Natural Sciences and Engineering Research Council of Canada) fellowship (S.A.-B.) in Canadian Government Laboratories Program.

## ACKNOWLEDGEMENTS

We would like to thank Camille Albouy, Fabien Leprieur and Arnaud Mosnier for fruitful discussion about generalized additive models and the numerous Department of Fisheries and Oceans (DFO) personnel for their contributions to data and sample collection, analysis and reporting. The authors would like to thank also the referees for their comments that improved the manuscript.

## REFERENCES

Alcaraz, M., Felipe, J., Grote, U., Arashkevich, E. and Nikishina, A. B. (2014) Life in a warming ocean: thermal thresholds and metabolic balance of arctic zooplankton. *J. Plankton Res.*, **36**, 3–10.

- Allouche, O., Tsoar, A. and Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.*, **43**, 1223–1232.
- Barry, S. C. and Welsh, A. H. (2002) Generalized additive modelling and zero inflated count data. *Ecol. Model.*, **157**, 179–188.
- Bartón, K. (2014) Package “MuMIn”—Multi-model inference. Version 1.10.5, 1–47.
- Basedow, S. L., Tande, K. S. and Stige, L. C. (2010) Habitat selection by a marine copepod during the productive season in the Subarctic. *Mar. Ecol. Prog. Ser.*, **416**, 165–178.
- Beaugrand, G. and Helaouët, P. (2008) Simple procedures to assess and compare the ecological niche of species. *Mar. Ecol. Prog. Ser.*, **363**, 29–37.
- Borchers, D. L., Buckland, S. T., Priede, I. G. and Ahmadi, S. (1997) Improving the precision of the daily egg production method using generalized additive models. *Can. J. Fish. Aquat. Sci.*, **54**, 2727–2742.
- Broms, C. and Melle, W. (2007) Seasonal development of *Calanus finmarchicus* in relation to phytoplankton bloom dynamics in the Norwegian Sea. *Deep Sea Res. Part II*, **54**, 2760–2775.
- Broms, C., Melle, W. and Kaartvedt, S. (2009) Oceanic distribution and life cycle of *Calanus* species in the Norwegian Sea and adjacent waters. *Deep Sea Res. Part II*, **56**, 1910–1921.
- Buren, A. D., Koen-Alonso, M., Pepin, P., Mowbray, F., Nakashima, B., Stenson, G., Ollerhead, N. and Montevocchi, W. A. (2014) Bottom-up regulation of Capelin, a keystone forage species. *PLoS ONE*, **9**, e87589.
- Carstensen, J., Weymann, A., Olszewska, A. and Kwasniewski, S. (2012) Effects of environmental conditions on the biomass of *Calanus* spp. in the Nordic Seas. *J. Plankton Res.*, **34**, 951–966.
- Chust, G., Castellani, C., Licandro, P., Ibaibarriaga, L., Sagarminaga, Y. and Irigoien, X. (2014) Are *Calanus* spp. shifting poleward in the North Atlantic? A habitat modelling approach. *ICES J. Mar. Sci.*, **71**, 241–253.
- Conover, R. J. (1988) Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. *Hydrobiologia*, **167/168**, 127–142.
- Cunningham, R. B. and Lindenmayer, D. B. (2005) Modelling count data of rare species: some statistical issues. *Ecology*, **86**, 1135–1142.
- Darnis, G. and Fortier, L. (2014) Temperature, food and the seasonal vertical migration of key copepods in the thermally stratified Amundsen Gulf (Beaufort Sea, Arctic Ocean). *J. Plankton Res.*, **36**, 1092–1108.
- Davis, C. S. (1987) Zooplankton life cycles. In Backus, R. H. and Bourne, D. W. (eds), *Georges Bank*. MIT Press, Cambridge, pp. 256–267.
- Dawe, E. G., Koen-Alonso, M., Chabot, D., Stansbury, D. and Mullaney, D. (2012) Trophic interactions between key predatory fishes and crustaceans: comparison of two Northwest Atlantic systems during a period of ecosystem change. *Mar. Ecol. Prog. Ser.*, **469**, 233–248.
- Freeman, E. A. and Moisen, G. (2008) PresenceAbsence: an R package for presence-absence model analysis. *J. Stat. Softw.*, **23**, 1–31.
- Gislason, A., Astthorsson, O. S., Petursdottir, H., Gudfinnsson, H. and Bodvarsdottir, A. R. (2000) Life cycle of *Calanus finmarchicus* south of Iceland in relation to hydrography and chlorophyll *a*. *ICES J. Mar. Sci.*, **57**, 1619–1627.
- Grueber, C. E., Nakagawa, S., Laws, R. J. and Jamieson, I. G. (2011) Multimodel inferences in ecology and evolution: challenges and solutions. *J. Evol. Biol.*, **24**, 669–711.
- Head, E. J. H., Harris, L. R. and Campbell, R. W. (2000) Investigations on the ecology of *Calanus* spp. in the Labrador Sea. I. Relationship

- between the phytoplankton bloom and reproduction and development of *Calanus finmarchicus* in spring. *Mar. Ecol. Prog. Ser.*, **193**, 53–73.
- Head, E. J. H., Harris, L. R. and Petrie, B. (1999) Distribution of *Calanus* spp. on and around the Nova Scotia Shelf in April: evidence for an offshore source of *Calanus finmarchicus* to the central and western regions. *Can. J. Fish. Aquat. Sci.*, **56**, 2463–2476.
- Head, E. J. H., Harris, L. R. and Yashayaev, I. (2003) Distribution of *Calanus* spp. and other mesozooplankton in the Labrador Sea in relation to hydrography in spring and summer 1995–2000. *Prog. Oceanogr.*, **59**, 1–30.
- Head, E. J. H., Melle, W., Pepin, P., Bagoien, E. and Broms, C. (2013) On the ecology of *Calanus finmarchicus* in the Subarctic North Atlantic: a comparison of population dynamics and environmental conditions in areas of the Labrador Sea–Labrador/Newfoundland Shelf and Norwegian Sea Atlantic and Coastal Waters. *Prog. Oceanogr.*, **114**, 46–63.
- Head, E. J. H. and Pepin, P. (2007) Variations in overwintering depth distributions of *Calanus finmarchicus* in the slope waters of the NW Atlantic Continental Shelf and the Labrador Sea. *J. Northw. Atl. Fish. Sci.*, **39**, 49–69.
- Head, E. J. H. and Pepin, P. (2010) Spatial and inter-decadal variability in plankton abundance and composition in the Northwest Atlantic (1958–2006). *J. Plankton Res.*, **32**, 1633–1648.
- Head, E. J. H. and Sameoto, D. D. (2007) Inter-decadal variability in zooplankton and phytoplankton abundance on the Newfoundland and Scotian shelves. *Deep-Sea Res. Part II*, **54**, 2686–2701.
- Heinänen, S., Rönkä, M. and Von Numers, M. (2008) Modelling the occurrence and abundance of a colonial species, the arctic tern *Sterna paradisaea* in the archipelago of SW Finland. *Ecography*, **31**, 601–611.
- Helaouët, P. and Beaugrand, G. (2007) Macroecology of *Calanus finmarchicus* and *C. helgolandicus* in the North Atlantic Ocean and adjacent seas. *Mar. Ecol. Prog. Ser.*, **345**, 147–165.
- Helaouët, P., Beaugrand, G. and Edwards, M. (2013) Understanding long-term changes in species abundance using a niche-based approach. *PLoS ONE*, **8**, e79186.
- Helaouët, P., Beaugrand, G. and Reid, P. C. (2011) Macrophysiology of *Calanus finmarchicus* in the North Atlantic Ocean. *Prog. Oceanogr.*, **91**, 217–228.
- Helaouët, P., Beaugrand, G. and Reygondeau, G. (2015) Reliability of spatial and temporal patterns of *C. finmarchicus* inferred from the CPR survey. *J. Mar. Syst.*, **153**, 18–24.
- Herman, A. W., Sameoto, D. D., Shunlian, C., Mitchell, M. R., Petrie, B. and Cochrane, N. (1991) Sources of zooplankton on the Nova-Scotia Shelf and their aggregations within deep-shelf basins. *Cont. Shelf Res.*, **11**, 211–238.
- Hinder, S. L., Gravenor, M. B., Edwards, M., Ostle, C., Bodger, O. G., Lee, P. L. M., Walne, A. W. and Hays, G. C. (2014) Multi-decadal range changes vs. thermal adaptation for north east Atlantic oceanic copepods in the face of climate change. *Glob. Change Biol.*, **20**, 140–146.
- Hirche, H.-J. (2013) Long-term experiments on lifespan, reproduction activity and timing of reproduction in the Arctic copepod *Calanus hyperboreus*. *Mar. Biol.*, **160**, 2469–2481.
- Hirche, H.-J. and Kattner, G. (1993) Egg production and lipid content of *Calanus glacialis* in spring: indication of a food-dependent and food-independent reproductive mode. *Mar. Biol.*, **117**, 15–22.
- Hutchinson, G. E. (1957) Population studies—animal ecology and demography—concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.*, **22**, 415–427.
- Irigoien, X. and Harris, R. P. (2006) Comparative population structure, abundance and vertical distribution of six copepod species in the North Atlantic: evidence for intraguild predation? *Mar. Biol. Res.*, **2**, 276–290.
- Johnson, C., Runge, J. A., Curtis, A., Durbin, E. G., Hare, J. A., Incze, L. S., Link, J. S., Melvin, G. D. et al. (2011) Biodiversity and ecosystem function in the Gulf of Maine: pattern and role of zooplankton and pelagic nekton. *PLoS ONE*, **6**, e16491.
- Jónasdóttir, S. H. and Koski, M. (2011) Biological processes in the North Sea: a comparison of *Calanus helgolandicus* and *Calanus finmarchicus* vertical distribution and production. *J. Plankton Res.*, **33**, 85–103.
- Kane, J. (2005) The demography of *Calanus finmarchicus* (Copepoda: Calanoida) in the Middle Atlantic Bight, USA, 1977–2001. *J. Plankton Res.*, **27**, 401–414.
- Kosobokova, K. N. (1998) New data on the life cycle of *Calanus glacialis* in the White Sea (based on seasonal observations of its genital system development). *Oceanology*, **38**, 347–355.
- Kosobokova, K. N. (1999) The reproductive cycle and life history of the Arctic copepod *Calanus glacialis* in the White Sea. *Polar Biol.*, **22**, 254–263.
- Labat, J. P., Gasparini, S., Mousseau, L., Prieur, L., Boutoute, M. and Mayzaud, P. (2009) Mesoscale distribution of zooplankton biomass in the northeast Atlantic Ocean determined with an Optical Plankton Counter: relationships with environmental structures. *Deep-Sea Res. Part I*, **56**, 1742–1756.
- Longhurst, A. (1998) *Ecological Geography of the Sea*. Academic Press, London.
- Maps, F., Plourde, S. and Zakardjian, B. (2010) The control of dormancy by the lipid metabolism in *Calanus finmarchicus*: a population model test. *Mar. Ecol. Prog. Ser.*, **403**, 165–180.
- Maps, F., Record, N. R. and Pershing, A. J. (2014) A metabolic approach to dormancy in pelagic copepods helps explaining inter- and intra-specific variability in life history strategies. *J. Plankton Res.*, **36**, 18–30.
- McGinty, N., Power, A. M. and Johnson, M. P. (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.
- Meise, C. J. and O'Reilly, J. E. (1996) Spatial and seasonal patterns in abundance and age-composition of *Calanus finmarchicus* in the Gulf of Maine and on Georges Bank: 1977–1987. *Deep-Sea Res. Part II*, **43**, 1473–1501.
- Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., Pierson, J., Janasdóttir, S. et al. (2014) The North Atlantic Ocean as habitat for *Calanus finmarchicus*: environmental factors and life history traits. *Prog. Oceanogr.*, **129**, 244–284.
- Mitchell, M. R., Harrison, G., Pauley, K., Gagné, A., Maillet, G. and Strain, P. M. (2002) Atlantic zone monitoring program sampling protocol. *Canadian Data Report of Hydrography and Ocean Sciences*, p. 23.
- Møller, E. F., Maar, M., Jónasdóttir, S. H., Nielsen, T. G. and Tønnesson, K. (2012) The effect of changes in temperature and food on the development of *Calanus finmarchicus* and *Calanus helgolandicus*. *Limnol. Oceanogr.*, **57**, 221–220.
- Morata, N. and Søreide, J. E. (2015) Effect of light and food on the metabolism of the Arctic copepod *Calanus glacialis*. *Polar Biol.*, **38**, 67–73.
- Morissette, L., Castonguay, M., Savenkoff, C., Swain, D. P., Chabot, D., Bourdages, H., Hammill, M. O. and Hanson, J. M. (2009) Contrasting changes between the northern and southern Gulf of

- St. Lawrence ecosystems associated with collapse of groundfish stocks. *Deep-Sea Res. Part II*, **56**, 2117–2131.
- Niehoff, B. and Hirche, H.-J. (2005) Reproduction of *Calanus glacialis* in the Lurefjord (western Norway): indication for temperature-induced female dormancy. *Mar. Ecol. Prog. Ser.*, **285**, 107–115.
- Papastamatiou, Y. P., Meyer, C. G., Carvalho, F., Dale, J. J., Hutchinson, M. R. and Holland, K. N. (2013) Telemetry and random-walk models reveal complex patterns of partial migration in a large marine predator. *Ecology*, **94**, 2595–2606.
- Pasternak, A. F., Arashkevich, E. G., Grothe, U., Nikishina, A. B. and Solov'yev, K. A. (2013) Different effects of increased water temperature on egg production of *Calanus finmarchicus* and *C. glacialis*. *Mar. Biol.*, **53**, 547–553.
- Pepin, P., Colbourne, E. and Maillet, G. (2011) Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador Shelf. *Prog. Oceanogr.*, **91**, 273–285.
- Pepin, P., Han, G. Q. and Head, E. J. (2013) Modelling the dispersal of *Calanus finmarchicus* on the Newfoundland Shelf: implications for the analysis of population dynamics from a high frequency monitoring site. *Fish. Oceanogr.*, **22**, 371–387.
- Pepin, P., Johnson, C. L., Harvey, M., Casault, B., Chassé, J., Colbourne, E. B., Galbraith, P. S., Hebert, D. et al. (2015) A multi-variate evaluation of environmental effects on zooplankton community structure in the western North Atlantic. *Prog. Oceanogr.*, <http://dx.doi.org/10.1016/j.pocean.2015.01.017>.
- Pepin, P., Petrie, B., Theriault, J.-C., Narayanan, S., Harrison, W. G., Frank, K. T., Chassé, J., Colbourne, E. B. et al. (2005) The Atlantic Zone Monitoring Program (AZMP): review of 1998–2003. *Canadian Technical Report of Hydrography and Ocean Science*, Vol. **242**, pp. v+87.
- Planque, B., Loots, C., Petitgas, P., Lindstrom, U. and Vaz, S. (2011) Understanding what controls the spatial distribution of fish population using a multi-model approach. *Fish. Oceanogr.*, **20**, 1–17.
- Plourde, S., Campbell, R. G., Ashjian, C. J. and Stockwell, D. A. (2005) Seasonal and regional patterns in egg production of *Calanus glacialis/marsshallae* in the Chukchi and Beaufort Seas during spring and summer, 2002. *Deep-Sea Res. Part II—Top. Stud. Oceanogr.*, **52**, 3411–3426.
- Plourde, S., Joly, P., Runge, J. A., Dodson, J. and Zakardjian, B. (2003) Life cycle of *Calanus hyperboreus* in the lower St. Lawrence Estuary and its relationship to local environmental conditions. *Mar. Ecol. Prog. Ser.*, **255**, 219–233.
- Plourde, S., Joly, P., Runge, J. A., Zakardjian, B. and Dodson, J. J. (2001) Life cycle of *Calanus finmarchicus* in the Lower St. Lawrence Estuary: the imprint of circulation and late timing of the spring phytoplankton bloom. *Can. J. Fish. Aquat. Sci.*, **58**, 647–658.
- Plourde, S., Pepin, P. and Head, E. J. H. (2009) Long-term seasonal and spatial patterns in mortality and survival of *Calanus finmarchicus* across the Atlantic Zone Monitoring Programme region, Northwest Atlantic. *ICES J. Mar. Sci.*, **66**, 1942–1958.
- Plourde, S., Starr, M., Devine, L., St-Pierre, J.-F., St-Amand, L., Joly, P. and Galbraith, P. S. (2014) Chemical and biological oceanographic conditions in the Estuary and Gulf of St. Lawrence during 2011 and 2012. *DFO Canadian Science Advisory Secretariat Research Document*, 2014/049, pp. v+46.
- R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>.
- Razouls, C., De Bovée, F., Kouwenberg, J. and Desreumaux, N. (2015) Diversité et répartition géographique chez les Copépodes planctoniques marins. <http://copepodes.obs-banyuls.fr>.
- Runge, J. A., Ji, R., Thomson, C. R. S., Record, N. R., Chen, C., Vandemark, D. C., Salisbury, J. E. and Maps, F. (2015) Persistence of *Calanus finmarchicus* in the western Gulf of Maine during recent extreme warming. *J. Plankton Res.*, **37**, 221–232.
- Runge, J. A., McLaren, I. A., Corkett, C. J., Bohrer, R. N. and Koslow, J. A. (1985) Molting rates and cohort development of *Calanus finmarchicus* and *C. glacialis* on the sea off southwest Nova Scotia. *Mar. Biol.*, **86**, 241–246.
- Sameoto, D. D. and Herman, A. W. (1992) Effect of the outflow from the Gulf of St-Lawrence on Nova-Scotia Shelf zooplankton. *Can. J. Fish. Aquat. Sci.*, **49**, 857–869.
- Theriault, J.-C., Petrie, B., Pepin, P., Gagnon, J., Gregory, D., Helbig, J., Herman, A., Lefavre, D. et al. (1998) Proposal for a Northwest Atlantic Zonal Monitoring Program. *Canadian Technical Reports of Hydrography and Ocean Sciences*, **194**, pp. vii+57.
- Williams, R. (1985) Vertical distribution of *Calanus finmarchicus* and *C. helgolandicus* in relation to the development of the seasonal thermocline in the Celtic Sea. *Mar. Biol.*, **86**, 145–149.
- Wintle, B. A., Elith, J. and Potts, J. M. (2005) Fauna habitat modelling and mapping: a review and case study in the Lower Junter Central Coast Region of NSW. *Aust. Ecol.*, **30**, 719–738.
- Wood, S. (2006) Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, **62**, 1025–1036.
- Wood, S. and Scheipl, F. (2014) Package “gamm4”: general additive mixed models using mgcv and lme4 version 0.2–3, 1–9.
- Yee, T. W. and Mitchell, N. D. (1991) Generalized additive models in plant ecology. *J. Veg. Sci.*, **2**, 587–602.