The following supplements accompany the article

Temperature effects on *Calanus finmarchicus* vary in space, time and between developmental stages

Kristina Øie Kvile[1,]*, Padmini Dalpadado[2], Emma Orlova[3,†], Nils C. Stenseth[1], Leif C. Stige[1]

[1]**Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway**

[2]**Institute of Marine Research, PO Box 1870 Nordnes, 5817 Bergen, Norway**

[3]**Knipovich Polar Research Institute of Marine Fisheries and Oceanography, Knipovich-St. 6, 183763 Murmansk, Russia**

*Corresponding author: k.o.kvile@ibv.uio.no

Marine Ecology Progress Series 517: 85–104 (2014)

Supplement 1. Supplementary modelling material

Further background on generalised additive models (GAMs) can be found in Hastie & Tibshirani (1990) and Wood (2006). Note that the smooth terms are constrained to have a mean of 0 in order for the model to be identifiable (Wood 2006). Random effects are treated as smooths by setting 'bs = re' when specifying the smooth in the mgcv library (Wood 2013). Smooth functions varying with a linear or factor variable are implemented using the 'by' operator within the smooth function in the mgcv library. We used a tensor product (setting 'te' instead of 's' when formulating the GAMs in the mgcv library) for the effect of geographical position because the isotropy in longitude and latitude measurements is reduced when we approach the poles.

Equations described in the main text (Eqs. 1-6 in Table 1)

Eq. (1): Spatial, seasonal and vertical variation

$$Z_{l,t,d} = \beta + s_1(j_t) + s_2(d_d) + te(x_l,y_l)l_f + s_3(Y) + \varepsilon_{l,t,d}$$

$Z_{l,t,d}$ is stage-specific abundance ($\log_e$ [n+1]) at location l, time t and depth d; β is the intercept; $s_1(j_t)$ and $s_2(d_d)$ are 1-dimensional smooth functions of, respectively, day-of-year and average sampling depth (cubic regression splines with maximally 3 degrees of freedom [df]); $te(x_l,y_l)$ is a 2-dimensional tensor product of longitude and latitude representing the effect of position (thin-plate regression splines with maximally 6 df each); l_f is an indicator variable of season (spring or summer) multiplied with the tensor product of longitude and latitude; $s_3(Y)$ is a random effect of year added to capture year-to-year variation not explained by seasonal or spatial variation; and $\varepsilon_{l,t,d}$ is a random error term.

Eq. (2): Interaction between position and day-of-year

$$Z_{l,t} = \beta + te(x_l,y_l,j_t) + \varepsilon_{l,t}$$

$Z_{l,t}$, β and $\varepsilon_{l,t}$ correspond to similar terms in Eq. (1), while $te(x_l,y_l,j_t)$ is an interaction term between position and day-of-year, formulated as a tensor product of a 2-dimensional smooth function of longitude and latitude (thin-plate regression spline with maximally 7 df) and a 1-dimensional smooth function of day-of-year (cubic regression spline with maximally 3 df). The position term (longitude and latitude) was formulated here as a 2-dimensional smooth (instead of 2 separate smooths) to avoid over-parameterisation of the model. We investigated the model for samples from the upper water layer only to facilitate comparison with the more complex temperature model (Eq. 6 in Table 1 of the main text), and to isolate seasonal variation in the upper waters from seasonal vertical migration.

Eq. (3): Temperature effect varying between stages

$$Z_{l,t,d} = \beta + s_1(j_t) + s_2(d_d) + te(x_l,y_l)l_f + s_3(T_{l,t,d}) + \varepsilon_{l,t,d}$$

$Z_{l,t,d}$, β, $s_1(j_t)$, $s_2(d_d)$, $te(x_l,y_l)l_f$ and $\varepsilon_{l,t,d}$ correspond to the first part of Eq. (1) and form the null-model core of the model formulation. $s_3(T_{l,t,d})$ is a 1-dimensional smooth function of local temperature anomaly (cubic regression spline with maximally 3 df).

Eq. (4): Seasonally varying temperature effect

$$Z_{l,t,d} = \beta + s_1(j_t) + s_2(d_d) + te(x_l,y_l)l_f + s_3(T_{l,t,d})l_f + \varepsilon_{l,t,d}$$

$Z_{l,t,d}$, β, $s_1(j_t)$, $s_2(d_d)$, $te(x_l,y_l)l_f$, $s_3(T_{l,t,d})$ and $\varepsilon_{l,t,d}$ correspond to similar terms in Eq. (3). l_f is an indicator variable of season (spring or summer) that is multiplied with the smooth function of local temperature anomaly ($s_3(T_{l,t,d})$).

Eq. (5) Spatially varying temperature effect

$$Z_{l,t} = \beta + s_1(j_t) + te_1(x_l,y_l)l_f + te_2(x_l,y_l)l_fT_{l,t} + \varepsilon_{l,t}$$

$Z_{l,t}$, β, $s_1(j_t)$, $te_1(x_l,y_l)l_f$ and $\varepsilon_{l,t}$ correspond to the first part of Eq. (1) and form the null-model core of the model formulation, but as the model is formulated for samples from the upper depth layer only, the smooth function of depth is removed. $te_2(x_l,y_l)l_fT_{l,t}$ is a 2-dimensional tensor product smooth of longitude and latitude (thin-plate regression splines with maximally 3 df each) that is multiplied with a factor variable of season (l_f) and with a linear temperature term ($T_{l,t}$).

Eq. (6): Spatially and temporally varying temperature effect

$$Z_{l,t} = \beta + te_1(x_l,y_l,j_t) + te_2(x_l,y_l,j_t)T_{l,t} + \varepsilon_{l,t}$$

$Z_{l,t}$, β, $te_1(x_l,y_l,j_t)l_f$ and $\varepsilon_{l,t}$ correspond to the first part of Eq. (2) and constitute the null-model core of the model formulation. $te_2(x_l,y_l,j_t)$ is a tensor product of a 2-dimensional smooth function of longitude and latitude (thin-plate spline with maximally 3 df) and a 1-dimensional smooth function of day-of-year (cubic regression spline with maximally 2 df). This second interaction term is multiplied by a linear term of temperature anomaly, $T_{l,t}$. The null-model base of the model formulation was changed compared to the previous model formulations (Eqs. 3–5) to assess whether adding a linear temperature term varying smoothly with position and day-of-year to a model already capturing the interaction between the variables would improve the model's predictive power. The degrees of freedom in the second interaction term had to be reduced to avoid over-parameterisation of the model.

Additional equations (Eqs. S1–S4)

Eq. (S1): Estimation of local temperature anomalies

$$T_{l,t,d} = \beta + te(x_l,y_l) + s(j_t)l_d + d_d + \varepsilon_{l,t,d}$$

$T_{l,t,d}$ is a local temperature estimate extracted from a numerical ocean model hindcast archive (Lien et al. 2013) at location l, time t and depth d (see main text for details); β is the intercept; $te(x_l,y_l)$ is a 2-dimensional tensor product of longitude and latitude (natural cubic regression splines with maximally 4 df each); $s(j_t)$ is a 1-dimensional smooth function of day of year (cubic regression spline with maximally 4 df), multiplied by l_d, a factor variable of depth layer (upper, middle or lower); d_d is a factor variable of depth layer (upper, middle or lower); and $\varepsilon_{l,t,d}$ is a random error term.

Eq. (S2): Indices of annual spring and summer stage-specific abundances

$$Z_{l,t,d} = f(Y) + s_1(j_t) + s_2(d_d) + te(x_l,y_l) + \varepsilon_{l,t,d}$$

Indices were constructed by extracting year-specific intercepts from the model, formulated separately for spring and summer samples. $Z_{l,t,d}$ is stage-specific abundance ($\log_e$ [n+1]) at location l, time t and depth d; $f(Y)$ is a year-specific intercept; $s_1(j_t)$ and $s_2(d_d)$ are 1-dimensional smooth functions of, respectively, day-of-year and average sampling depth (cubic regression splines with maximally 3 df); $te(x_l,y_l)$ is a 2-dimensional tensor product of longitude and latitude (thin-plate regression splines with maximally 6 df each); and $\varepsilon_{l,t,d}$ is a random error term.

Eq. (S3): Seasonally varying depth effect

$$Z_{l,t,d} = \beta + s_1(j_t) + s_2(d_d)l_f + te(x_l,y_l)l_f + s_3(Y) + \varepsilon_{l,t,d}$$

$Z_{l,t,d}$, β, $s_1(j_t)$, $s_2(d_d)$, $te(x_l,y_l)l_f$, $s_3(Y)$ and $\varepsilon_{l,t,d}$ correspond to similar terms in Eq. (1). l_f is an indicator variable of season (spring or summer) multiplied with the smooth function of depth.

Eq. (S4): Seasonally and vertically varying temperature effect

$$Z_{l,t,d} = \beta + s_1(j_t) + s_2(d_d) + te(x_l,y_l)l_f + s_3(T_{l,t,d})l_{f,d} + \varepsilon_{l,t,d}$$

$Z_{l,t,d}$, β, $s_1(j_t)$, $s_2(d_d)$, $te(x_l,y_l)l_f$, $s_3(T_{l,t,d})$ and $\varepsilon_{l,t,d}$ correspond to similar terms in Eq. (3). $l_{f,d}$ is an indicator variable of both season (spring or summer) and depth (upper, middle or lower) that is multiplied with the smooth function of local temperature anomaly ($s_3(T_{l,t,d})$). The temperature term is thus allowed to vary between all combinations of season and depth layer.

LITERATURE CITED

Hastie T, Tibshirani R (1990) Generalized additive models. Chapman & Hall, London

Lien VS, Gusdal Y, Albretsen J, Melsom A, Vikebø F (2013) Evaluation of a Nordic Seas 4 km numerical ocean model hindcast archive (SVIM), 1960-2011. Fisken Havet 7:1–80

Wood SN (2006) Generalized additive models: an introduction with R. Chapman & Hall/CRC, Boca Raton, FL

Wood SN (2013) mgcv: GAMs with GCV smoothness estimation and GAMMs by REML/ PQL. R package. Version 1.7–24. http://lojze.lugos.si/~darja/software/r/library/mgcv/html/mgcv-package.html

Supplement 2. Genuine cross-validation and model comparison

Models were compared with genuine cross-validation (GCV) by considering year as the sampling unit. This measure was chosen as data samples were not independent in space or time within years, possibly leading to positive auto-correlation in the model residuals. Using generalised CV or AIC in model selection would in this case likely favour over-parameterised models. GCV was calculated as the mean of the mean-squared predictive error from 1000 models formulated from datasets where data from 1 randomly chosen year were removed, and the observations from the removed year were predicted from a new model based on the reduced dataset. The GCV increases with high complexity and low predictive power, and models with lower GCV values are therefore better. Significance of model terms was assessed by comparing GCV and R^2 (the proportion of data variation explained by the model) with and without the model term in question, and by plotting the model term and assessing whether the 95% bootstrap intervals of the model effect differed from 0. The number of knots (i.e. the flexibility of the model terms / degrees of freedom) in model terms were also determined by minimising the GCV, and by generally keeping model complexity within ecologically reasonable limits.

Supplement 3. Additional figures and tables

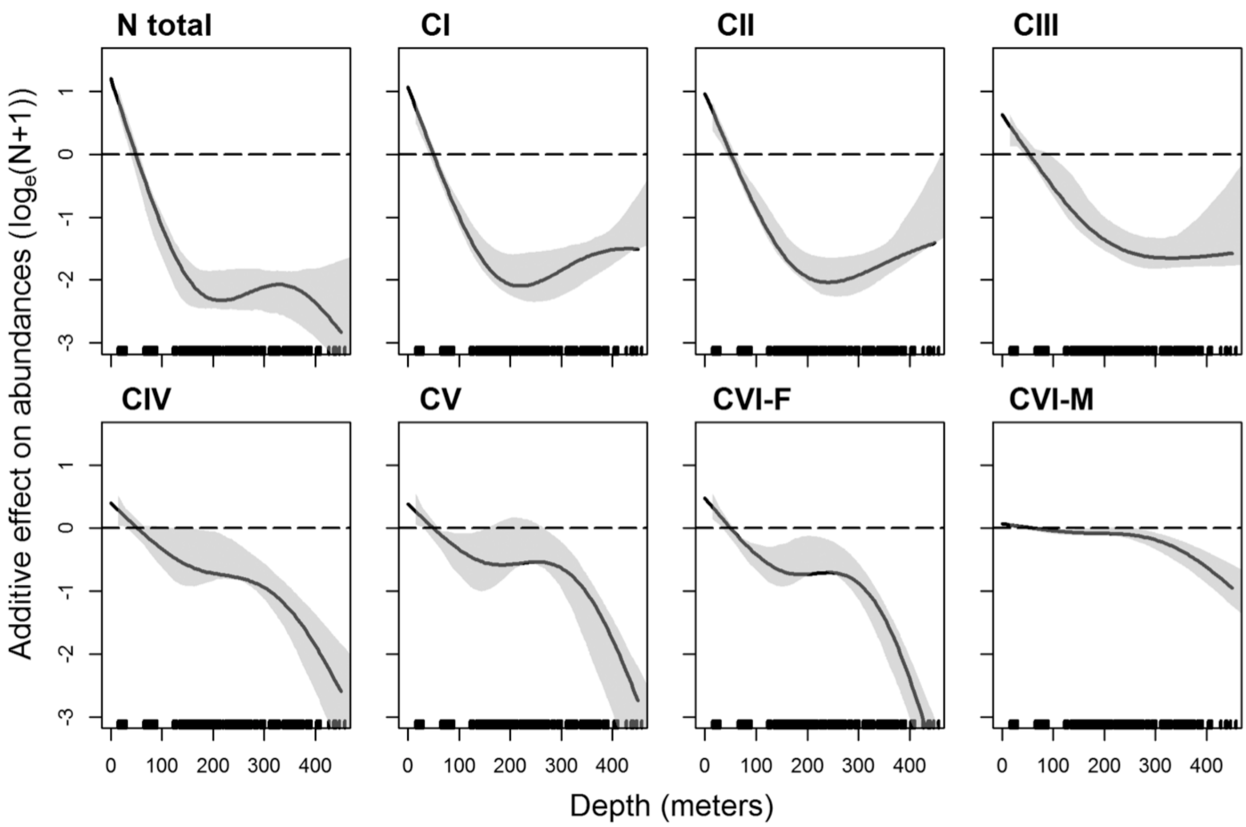


Fig. S1. Additive effect of depth on *Calanus finmarchicus* stage-specific abundances (log$_e$[n+1]) (see Eq. 1 in Table 1 of the main text and Supplement 1). Shaded area: 95% confidence interval from bootstrap procedure. Dashed line: 0 effect isoline. N total: total nauplii abundance, CI–CVI: stage-specific copepodite abundances, F: female, M: male

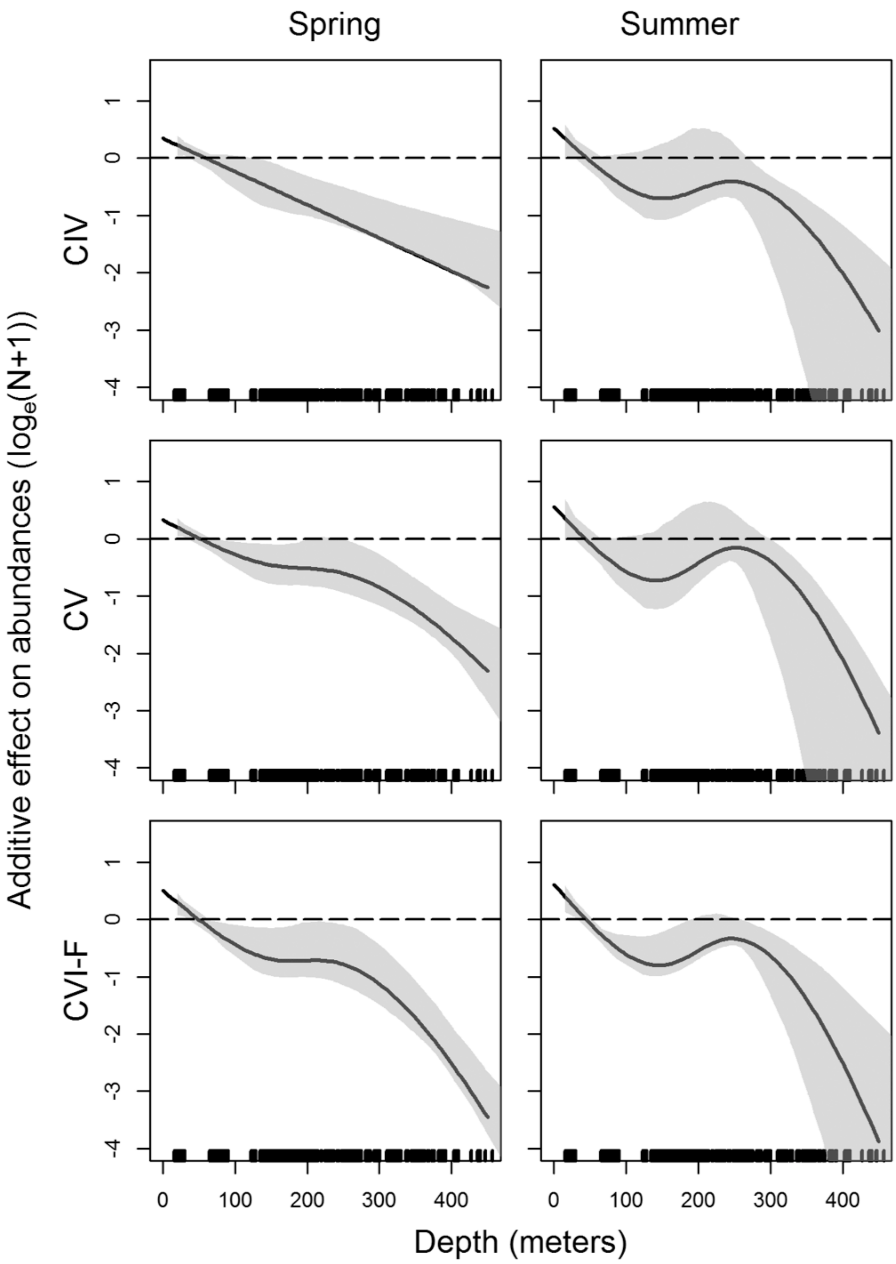
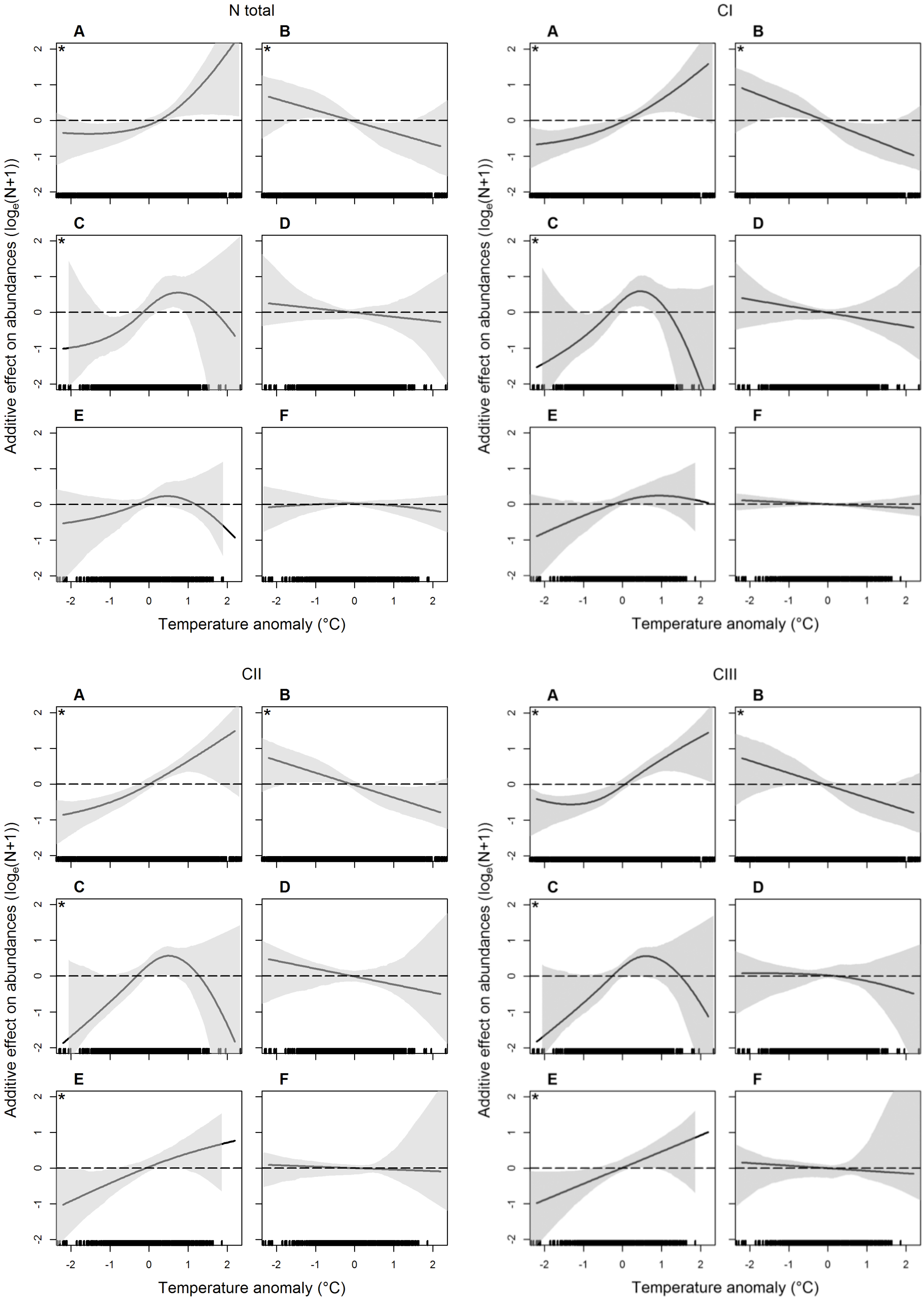


Fig. S2. Additive effect of depth on abundances ($\log_e$[n+1]) of *Calanus finmarchicus* copepodite stages CIV–CVI-F (female) in spring (left) and summer (right) (see Eq. S3 in Supplement 1). Shaded area: 95% confidence interval from bootstrap procedure. Dashed line: 0 effect isoline


N total
CI
CII
CIII
A
B
C
D
E
F
Additive effect on abundances ($log_e(N+1)$)
Temperature anomaly (°C)

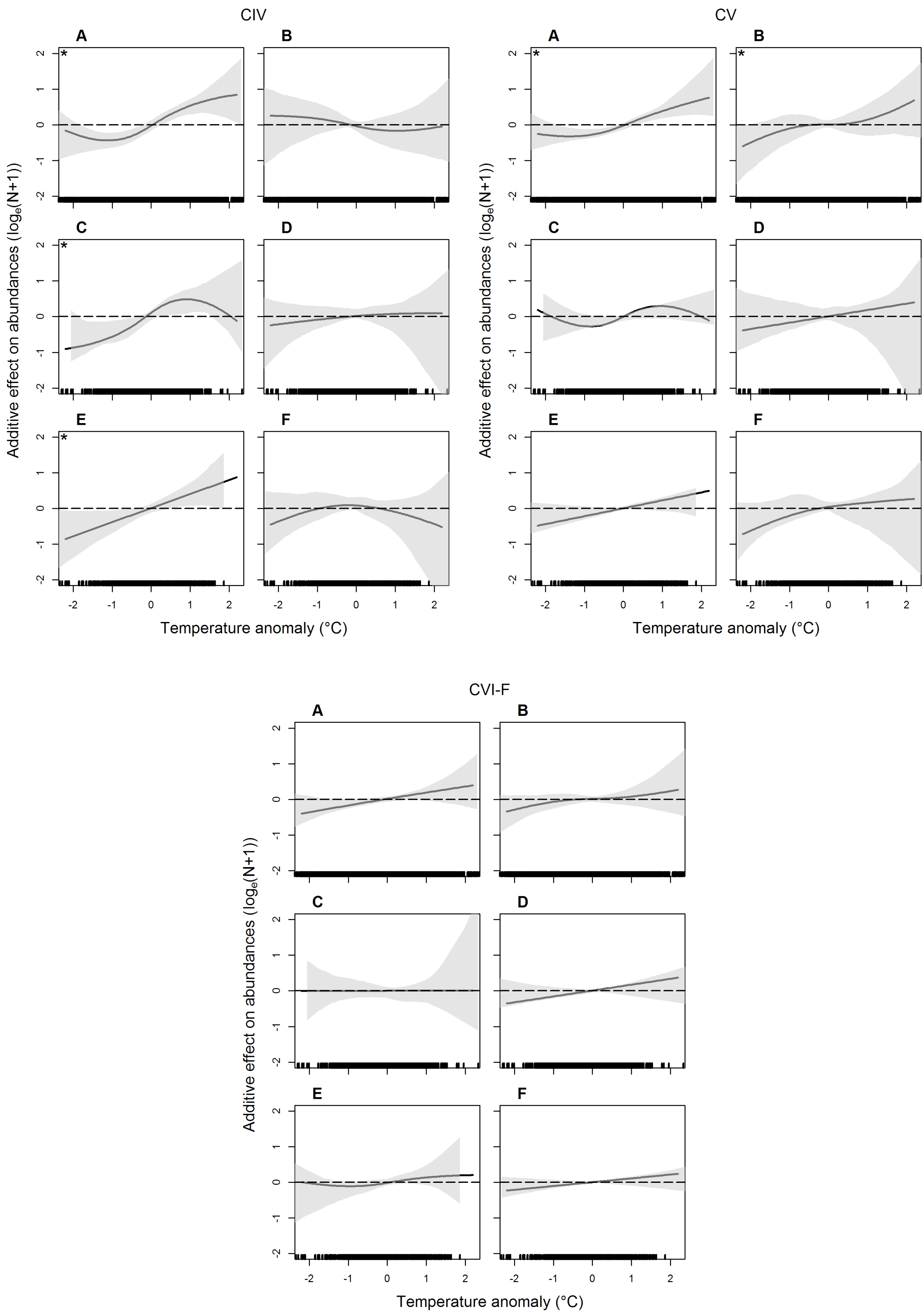


Fig. S3. Additive effect of local temperature anomalies on *Calanus finmarchicus* stage-specific abundances ($\log_e$[n+1]) compared across depth layers and seasons (Eq. S4 in Supplement 1). Six panels are displayed for each development stage: (A) upper water layer in spring; (B) upper water layer in summer; (C) middle water layer in spring; (D) middle water layer in summer; (E) lower water layer in spring; (F) lower water layer in summer. Shaded area: 95% confidence interval from bootstrap procedure. Dashed line: 0 effect isoline. Stars indicate a significant association, i.e. that the effect differs from 0 in parts of the covariates' range. N total: total nauplii abundance, CI–CVI: stage-specific copepodite abundance, F: female

Table S1. Importance of variables explaining general spatiotemporal variation in *Calanus finmarchicus* abundance: R^2 and genuine cross-validation (GCV) for models with all predictors (see Eq. 1 in Table 1 of the main text and Supplement 1, 'full model') and models with 1 term omitted. The difference in R^2 from the full model indicates the amount of data variation explained by the variable in question. The difference in GCV indicates the reduction in model predictive power when removing the variable. The most important variable(s) in terms of R^2 and GCV per developmental stage are shown in **bold**. N total: total nauplii abundance, CI–CVI: stage-specific copepodite abundance, F: female, M: male

	Full model		**– Position**		**– Day**		**– Depth**	
	R^2	**GCV**	**R^2**	**GCV**	**R^2**	**GCV**	**R^2**	**GCV**
N total	0.45	1.85	**0.28**	2.03	0.36	**3.88**	0.30	2.07
CI	0.45	1.67	**0.24**	1.92	0.39	**3.10**	0.31	1.89
CII	0.40	1.64	**0.18**	1.87	0.37	**2.10**	0.26	1.84
CIII	0.30	1.70	**0.10**	1.87	0.26	**1.88**	0.22	1.80
CIV	0.26	1.73	**0.14**	1.82	0.16	**3.81**	0.23	1.78
CV	0.37	1.62	**0.18**	1.79	0.22	**4.47**	0.34	1.67
CVI-F	0.50	1.27	**0.24**	**1.51**	0.49	1.27	0.43	1.38
CVI-M	0.42	0.48	**0.19**	0.55	0.39	**0.74**	0.40	0.49

Table S2. Spearman rank correlation coefficients (r_s) between *Calanus finmarchicus* stage-specific seasonal abundance indices (Eq. S2 in Supplement 1) and temperature indices from the Kola section and Skrova station. Significance level: *p < 0.05, **p < 0.01, ***p < 0.001. N total: total nauplii abundance, CI–CVI: stage-specific copepodite abundance, F: female

		Spring abundance						
	Temperature	N	CI	CII	CIII	CIV	CV	CVI–F
Kola	**Winter**	0.13	0.07	0.21	0.16	0.14	–0.03	–0.22
	Spring	0.2	0.2	0.34	0.29	0.32	0.15	–0.22
	Summer	0.18	0.26	0.39*	0.39*	0.45*	0.26	–0.16
Skrova	**Winter**	0.13	0.03	0.15	0.2	0.18	–0.04	–0.2
	Spring	0.04	–0.03	0.06	0.12	0.14	–0.07	–0.26
	Summer	0.18	0.23	0.35*	0.36*	0.41*	0.15	–0.1
		Summer abundance						
		N	CI	CII	CIII	CIV	CV	CVI–F
Kola	**Winter**	–0.21	–0.39*	–0.47*	–0.58**	–0.47*	–0.16	–0.12
	Spring	–0.29	–0.49*	–0.52**	–0.63**	–0.46*	–0.12	–0.08
	Summer	–0.35	–0.51*	–0.51*	–0.58**	–0.39*	–0.08	–0.04
Skrova	**Winter**	–0.01	–0.33	–0.39	–0.37	–0.28	0.04	0.01
	Spring	–0.13	–0.42*	–0.45*	–0.42*	–0.34	–0.02	–0.02
	Summer	–0.2	–0.39	–0.38	–0.39	–0.17	0.01	0.13