Oikos

OIK-05886

Pandori, L. L. M. and Sorte, C. J. B. 2018. The weakest link: sensitivity to climate extremes across life stages of marine invertebrates. – Oikos doi: 10.1111/oik.05886

Appendix 1

Search string in Web of Science

The search string below was used to search for papers documenting the effects of climate change stressors on multiple life stages of marine invertebrates. It contains sections searching for the topics of: (a) climate change stressors, (b) life stages, and (c) marine habitats. Resulting peer-reviewed manuscripts published until 1 January 2017 were reviewed for possible inclusion in the meta-analysis.

Query: TS = (climate change OR climatic change OR global change OR global warming OR temperature OR warm* OR heat* OR pH OR acidif* OR oxygen OR hypoxi* OR anoxi* OR CO2 OR carbon dioxide OR UV OR ultraviolet OR salinity OR desicc*) AND TS = (age OR life stage OR life history OR ontogen* OR juvenile* OR adult* OR larva* OR size) AND TS = (marine OR ocean OR saltwater OR intertidal OR subtidal OR mudflat OR estuarine OR saltmarsh OR pelagic OR benthic OR reef OR shore)

Appendix 2

Analysis of condition metric data

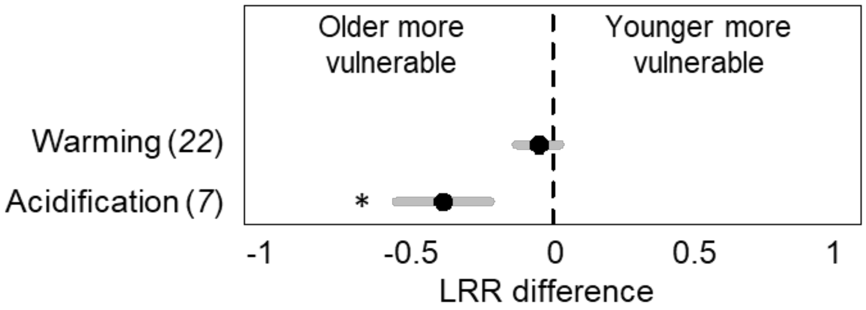


Figure A1. Condition metric (non-lethal) response differences between life stages marine invertebrates to climate change stressors in the paired analysis. Means ± bootstrap 95% CI's for effect size (LRR) differences. * = significant difference. Positive values indicate that younger life stages are more negatively affected, while negative values indicate that older life stages are more negatively affected by a given stressor. There were no data available that met our inclusion criteria documenting non-lethal responses to hypoxia.

Appendix 3

Meta data and R code

Data are available at <https://doi.org/10.15146/R3J09X>

R code is available at <https://github.com/llmpandori/Weakest-Link>

Appendix 4

Effect size metric comparison (LRR and Hedges d)

To calculate whether use of an unweighted or weighted metric would lead to differing conclusions, we calculated both for unpaired data, and have included the results of our investigation in the table below.

Hedges d = $\frac{\bar{Y}_1-\bar{Y}_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}} * J$, where J = $1 - \frac{3}{4(n_1+n_2-2)-1}$

(Hedges and Olkin 1985, Koricheva et al. 2013)

Table A1. Two effect size metrics (LRR and Hedges d) are presented for the unpaired analysis. However, due to the necessity of reported variance values for Hedges d, a weighted metric, sample sizes are smaller. We chose to use LRR as an effect size metric because it allows for the inclusion of more metadata in our analyses.

Stressor	Life Stag	LRR				Hedge's d			
		Mean	Upper CI	Lower CI	N	Mean	Upper CI	Lower CI	N
Warming	Adult	-0.6391784	-1.397973	0.10615627	30	-2.5005723	-9.7338342	5.02916711	20
	Juvenile	-1.2316316	-1.8893475	-0.5697793	24	-6.6465399	-11.126986	-2.0828222	16
	Larva	-1.7027215	-2.0970778	-1.3064244	87	-28.352806	-51.912039	-4.7788193	31
	Embryo	-1.8424027	-2.3342773	-1.3503483	72	-6.1020735	-15.219836	3.18050334	34
Acidification	Adult	-0.8259244	-1.2545809	-0.3942712	14	-0.0139202	-2.3536068	2.32570145	12
	Juvenile	-0.6388702	-1.2147155	-0.0591255	10	1.69076252	0.51873343	2.87300933	9
	Larva	-1.2745683	-2.1661599	-0.3788208	10	32.1035558	-8.8049825	72.264755	9
	Embryo	-1.2621322	-2.0097557	-0.5175927	4	ND	ND	ND	0
Hypoxia	Adult	-7.5417128	-10.813431	-4.3008723	9	13.6848739	4.33086026	23.1248983	2
	Juvenile	-8.5161987	-10.790868	-6.2603468	14	16.73249	-9.0381469	42.6796493	4
	Larva	-5.0461823	-7.1974291	-2.8946723	15	4.14640982	2.43733446	5.85509891	9
	Embryo	ND	ND	ND	0	ND	ND	ND	0

Appendix 5

Paired analysis with study and species as units of replication

As an additional test of the robustness of our conclusions in the paired analysis, we took the mean effect size for each species included in our analyses, and constructed bootstrap 95% confidence intervals (CIs) for each stressor in the paired analysis of lethal threshold data.

Table A2. There are no differences in direction or significance of paired analysis results if study or species is the unit of replication in our analyses. We used study as the unit of replication in our analyses, because it allowed for larger sample sizes.

Stressor	LRR - Study as Unit of Replication				LRR - Species as Unit of Replication			
	Mean	Upper CI	Lower CI	N	Mean	Upper CI	Lower CI	N
Warming	2.016065154	1.22738432	2.794331208	75	2.29450518	0.81740286	3.7784278	22
Acidification	0.0167788	-0.00848229	0.042075327	31	0.03844457	-0.02773831	0.1036387	6
Hypoxia	-1.425709599	-3.026975534	0.197330193	17	-0.17260225	-2.41873332	2.0535003	6

Appendix 6

Linear mixed-effects REML models using the rma.mv function in the Metafor package in R

Lethal threshold responses to warming – unpaired analysis

Model: We examined the effects of life stage, habitat (intertidal/subtidal), treatment magnitude and treatment duration on responses to warming (LRR), accounting for the random effect of study.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage + habitat + tmt_mag + tmt_time, random = ~1|study, data = temp)

Test for residual heterogeneity:

$Q_{E\ 206} = 814.50$, p = <0.0001

$Q_{M\ 6} = 131.67$, p = <0.0001

Model results:

Factor	Z-value	p-value
Life stage	3.94	<0.001
Habitat	-0.66	0.51
Treatment magnitude	-11.06	<0.001
Treatment duration	-1.73	0.08

Lethal threshold responses to acidification – unpaired analysis

Model: We examined the effects of life stage, habitat (intertidal/subtidal), treatment magnitude and treatment duration on responses to acidification (LRR), accounting for the random effect of study.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + + habitat + tmt_mag + tmt_time, random = ~1|study, data = ph)

Test for residual heterogeneity:

$Q_{E\ 31} = 30.29$, $p = 0.50$

$Q_{M\ 4} = 8.63$, $p = 0.07$

Model results:

Factor	Z-value	p-value
Life stage	0.84	0.40
Habitat	-0.88	0.38
Treatment magnitude	3.09	0.002
Treatment duration	-1.45	0.15

Lethal threshold responses to hypoxia – unpaired analysis

Model: We examined the effects of life stage and habitat (intertidal/subtidal) on effect size (LRR), accounting for the random effect of study.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + habitat, random = ~ 1|study, data = doxleth)

Test for residual heterogeneity:

$Q_{E\ 14} = 191.88$, $p < 0.0001$

$Q_{M\ 2} = 0.54$, $p = 0.76$

Model results:

Factor	Z-value	p-value
Life stage	0.14	0.89
Habitat	-0.58	0.56

Lethal responses to warming – paired analysis

Model: We examined the effects of life stage pairs, habitat (intertidal/subtidal), treatment magnitude, and treatment duration on the difference in effect sizes (LRR difference), accounting for the random effect of study.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + habitat + tmt_mag + tmt_time, random = ~1|study, data = tempxleth)

Test for residual heterogeneity:

$Q_{E\ 60} = 749.26$, $p < 0.0001$

$Q_{M\ 3} = 8.39$, $p = 0.04$

Model results:

Factor	Z-value	p-value
Life stage pair	3.22	0.001
Habitat	-2.64	0.01
Treatment magnitude	0.67	0.50
Treatment duration	0.78	0.43

Lethal responses to acidification – paired analysis

Model: We examined the effects of life stage pairs, habitat (intertidal/subtidal), treatment magnitude and treatment duration on the difference in effect sizes (LRR difference), accounting for the random effect of study.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + habitat + tmt_mag + tmt_time, random = ~ 1|study, data = phxleth)

Test for residual heterogeneity:

$Q_{E\ 26} = 0.15$, p = 1

$Q_{M\ 4} = 0.01$, p = 1

Model results:

Factor	Z-value	p-value
Life stage pair	-0.08	0.96
Habitat	-0.01	0.99
Treatment magnitude	-0.04	0.96
Treatment duration	-0.05	0.96

Lethal responses to hypoxia – paired analysis

Model: We examined the effects of life stage pairs, habitat (intertidal/subtidal), and phylum on the difference in effect sizes (LRR difference), accounting for the random effect of study.

R Code: rma.mv(yi = lrrdif, V = Ones, mods = ~ old_young2 + int_sub2 + phylum2, random = ~1|study, data = doxleth.data)

Test for residual heterogeneity:

$Q_{E\ 14} = 191.88$, $p < 0.0001$

$Q_{M\ 2} = 0.54$, p = 0.76

Model results:

Factor	Z-value	p-value
Life stage pair	-0.19	0.85
Habitat	0.14	0.89

Appendix 7

Accounting for phylogenetic differences and latitudinal origins

Accounting for differences in latitude of origin

Since all papers did not report the location of source populations for laboratory rearing or location of field collections, not all studies can be used in an analysis of the relationship between latitude and effect size. Our results show there is no significant relationship between latitude and effect size.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + lat2, random = ~1|study, data = tempxleth)

Test for residual heterogeneity:

$Q_{E\ 65} = 700.00$, $p < 0.0001$

$Q_{M\ 2} = 9.47$, $p = 0.01$

Model results:

Factor	Z-value	p-value
Life stage pair	3.07	0.002
Latitude	-0.20	0.84

Accounting for differences in phylogeny

To test for phylogenetic differences among study species, we ran a second model on paired data, with genus substituted for study as a random effect. Results are presented below. There were no significant differences in result magnitude or direction between the analyses.

Model: We examined the effects of life stage pairs, habitat (intertidal/subtidal), treatment magnitude, and treatment duration on the difference in effect sizes (LRR difference), accounting for the random effect of genus.

R Code: rma.mv(yi = lrr, V = Ones, mods = ~ lifestage2 + habitat + tmt_mag + tmt_time, random = ~1|genus2, data = tempxleth2)

Test for residual heterogeneity:

$Q_{E\ 52} = 585.64$, $p < 0.0001$

$Q_{M\ 4} = 8.39$, $p = 0.08$

Model results:

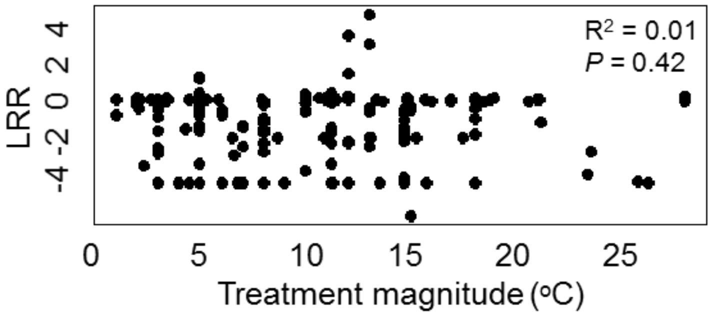
Factor	Z-value	p-value
Life stage pair	1.98	0.05
Habitat	-2.09	0.04
Treatment magnitude	-0.71	0.48
Treatment duration	0.16	0.87

Appendix 8

Regressions showing relationships between LRR and treatment magnitude

In our mixed-effects models, we found a significant effect of treatment magnitude on LRR for warming and acidification in the unpaired analysis. In order to investigate these trends, we conducted linear regressions between LRR and treatment magnitude for both stressors. Our results demonstrate a significant linear relationship for acidification, but not for warming.

(**a**) Warming



(**b**) Acidification

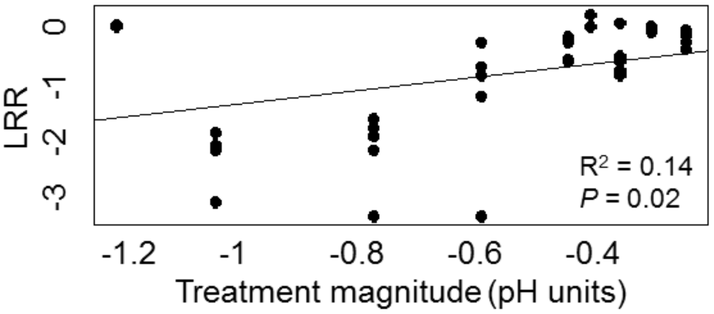


Figure S3. There is a statistically significant positive linear relationship between treatment magnitude and effect size (LRR) among acidification studies, but not for warming studies. Each point represents a unique study in our analysis. R^2 and P-values on figure panels were generated from linear regressions.

Appendix 9

Fail-safe numbers comparison

Two methods of calculating Fail-safe numbers were used in our analyses and are presented below. We utilized Rosenthal's (1979) criterion, where data were considered robust to sampling bias if the calculated fail-safe number was greater than 5N + 10 (where N is the number of observations utilized to construct bootstrapped confidence intervals).

Table A3. Fail-safe numbers for CI's in the unpaired (a) and paired (b) analyses indicate the number of studies with an effect size of 0 necessary to make a confidence interval non-significant. Numbers are reported for all significant values. In instances where CI's overlap 0, NA is reported. If no data were available, ND is reported. Two methods within the Metafor package are reported: Rosenthal's method and Rosenberg's method. Gray boxes indicate where failsafe numbers are below the 5N + 10 threshold of robustness (Rosenthal 1979). In the unpaired analysis, fail-safe numbers for warming and hypoxia, but not acidification, are robust to sampling bias. In the unpaired analysis, all fail-safe numbers are robust to sampling bias. These results are consistent across calculation methods.

(**a**) Unpaired analysis

Stressor	Life Stage	Fail-safe Numbers		
		Rosenthal	Rosenberg	5N + 10 Criterion
Warming	Adult	NA	NA	160
	Juvenile	299	204	130
	Larva	8024	5626	445
	Embryo	6432	4509	370
Acidification	Adult	36	21	80
	Juvenile	6	1	60
	Larva	51	33	60
	Embryo	6	3	30
Hypoxia	Adult	1694	1191	55
	Juvenile	5241	3687	80
	Larva	2103	1477	85
	Embryo	ND	ND	ND

(**b**) Paired analysis

Stressor	Fail-safe Numbers		
	Rosenthal	Rosenberg	5N + 10 Criterion
Warming	8376	5877	385
Acidification	NA	NA	165
Hypoxia	201	136	95