



META-ANALYSIS

## Strengthening confidence in climate change impact science

Mary I. O'Connor<sup>1,2,\*†</sup>, Johnna M. Holding<sup>3†</sup>, Carrie V. Kappel<sup>2</sup>, Carlos M. Duarte<sup>3,4</sup>, Keith Brander<sup>5</sup>, Christopher J. Brown<sup>6,7,8</sup>, John F. Bruno<sup>9</sup>, Lauren Buckley<sup>9</sup>, Michael T. Burrows<sup>10</sup>, Benjamin S. Halpern<sup>2,11,12</sup>, Wolfgang Kiessling<sup>13</sup>, Pippa Moore<sup>14,15</sup>, John M. Pandolfi<sup>16</sup>, Camille Parmesan<sup>17,18</sup>, Elvira S. Poloczanska<sup>7</sup>, David S. Schoeman<sup>19,20</sup>, William J. Sydeman<sup>21</sup> and Anthony J. Richardson<sup>7,22</sup>

<sup>1</sup>Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada V6T 1Z4, <sup>2</sup>National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, CA 93101, USA, <sup>3</sup>Department of Global Change Research, Instituto Mediterráneo de Estudios Avanzados, IMEDEA (UIB-CSIC), 07190 Esporles, Mallorca, Spain, <sup>4</sup>The UWA Oceans Institute and School of Plant Biology, University of Western Australia, 35 Stirling Highway, Crawley 6009, Australia, <sup>5</sup>Centre for Ocean Life, DTU Aqua, Charlottenlund Slot, 2920 Charlottenlund, Denmark, <sup>6</sup>School of Biological Sciences, The University of Queensland, St Lucia, Qld 4072, Australia, <sup>7</sup>Climate Adaptation Flagship, CSIRO Marine and Atmospheric Research, Ecosciences Precinct, Brisbane, Qld 4102, Australia, <sup>8</sup>Global Change Institute, The University of Queensland, St Lucia, Qld 4072, Australia, <sup>9</sup>Department of Biology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, <sup>10</sup>Scottish Association for Marine Science, Scottish Marine Institute, Oban, Argyll PA37 1QA, UK, <sup>11</sup>Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA, <sup>12</sup>Imperial College London, Department of Life Sciences, Silwood Park, Berkshire SL5 7PY, UK, <sup>13</sup>GeoZentrum Nordbayern, Paläoumwelt, Universität Erlangen-Nürnberg, Loewenichstrasse 28, 91054 Erlangen, Germany, <sup>14</sup>Centre for Marine Ecosystems Research, Edith Cowan University, Perth, WA 6027, Australia, <sup>15</sup>Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK, <sup>16</sup>Australian Research Council Centre of Excellence for Coral Reef Studies, School of Biological Sciences, The University of Queensland, St Lucia, Qld 4072, Australia, <sup>17</sup>Plymouth University, Drake Circus, Plymouth, Devon PL4 8AA, UK, <sup>18</sup>Section of Integrative Biology, University of Texas, Austin, TX 78712, USA, <sup>19</sup>Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore, DC, Qld, Australia, <sup>20</sup>Department of Zoology, Nelson Mandela Metropolitan University, PO Box 77000, Port Elizabeth, 6031, South Africa, <sup>21</sup>Farallon Institute for Advanced Ecosystem Research, 101 H Street, Suite Q, Petaluma, CA 94952, USA, <sup>22</sup>Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, University of Queensland, St Lucia, Qld 4072, Australia

\*Correspondence: Mary I. O'Connor, Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada V6T 1Z4. E-mail: [occonnor@zoology.ubc.ca](mailto:occonnor@zoology.ubc.ca)

†These authors contributed equally.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

### ABSTRACT

**Aim** To assess confidence in conclusions about climate-driven biological change through time, and identify approaches for strengthening confidence scientific conclusions about ecological impacts of climate change.

**Location** Global.

**Methods** We outlined a framework for strengthening confidence in inferences drawn from biological climate impact studies through the systematic integration of prior expectations, long-term data and quantitative statistical procedures. We then developed a numerical confidence index ( $C_{\text{index}}$ ) and used it to evaluate current practices in 208 studies of marine climate impacts comprising 1735 biological time series.

**Results** Confidence scores for inferred climate impacts varied widely from 1 to 16 (very low to high confidence). Approximately 35% of analyses were not associated with clearly stated prior expectations and 65% of analyses did not test putative non-climate drivers of biological change. Among the highest-scoring studies, 91% tested prior expectations, 86% formulated expectations for alternative drivers but only 63% statistically tested them. Higher confidence scores observed in studies that did not detect a change or tracked multiple species suggest publication bias favouring impact studies that are consistent with climate change. The number of time series showing climate impacts was a poor predictor of average confidence scores for a given group, reinforcing that vote-counting methodology is not appropriate for determining overall confidence in inferences.

**Main conclusions** Climate impacts research is expected to attribute biological change to climate change with measurable confidence. Studies with long-term, high-resolution data, appropriate statistics and tests of alternative drivers earn higher  $C_{\text{index}}$  scores, suggesting these should be given greater weight in impact assessments. Together with our proposed framework, the results of our  $C_{\text{index}}$  analysis indicate how the science of detecting and attributing biological impacts to climate change can be strengthened through the use of evidence-based prior expectations and thorough statistical analyses, even when data are limited, maximizing the impact of the diverse and growing climate change ecology literature.

### Keywords

**Abundance, attribution, climate change, distribution, hypothesis, impacts, marine, phenology, scientific method, time series.**

## INTRODUCTION

Increasing evidence that climate change has altered biological systems has spurred social, political and scientific concern (Parmesan, 2006; Rosenzweig *et al.*, 2008; Garcia *et al.*, 2014). Management decisions, future projections and scientific advancement require clear attribution of observed biological change to a suite of natural and anthropogenic pressures, including climate change (Hegerl *et al.*, 2010). Yet assessing confidence in the attribution of responses to climate change is difficult without generally accepted procedures for inferring climate impacts. As demand for reliable impact assessments grows and empirical research proliferates, a framework for assessing confidence in the purported attribution of biological change to climate change is urgently needed.

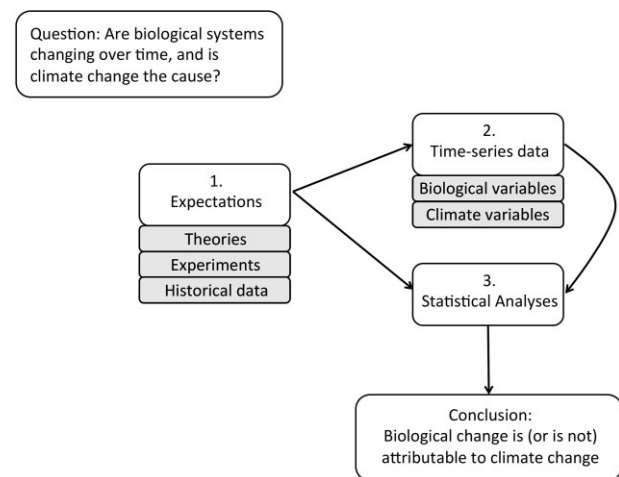
Attribution is the process of detecting biological change and inferring which pressures are the most likely causes of the change (Hegerl *et al.*, 2007, 2010). Inference about specific impacts of climate change as reported in individual studies often requires expert knowledge and a mechanistic understanding of how the system in question operates, which must then be conveyed to non-expert and non-specialist readers alike. There are a variety of approaches for causal attribution and inference in observational and time-series studies, and some convey stronger support for conclusions than others (Hegerl *et al.*, 2007, 2010; Morgan *et al.*, 2009). Some attribution methods allow for statistical confidence assessments (Hegerl *et al.*, 2010); however, simpler attribution methods are left with informal confidence assessments that are often based on 'expert opinion' and likelihood statements accompanied by probability assignments (*sensu* Hegerl *et al.*, 2007, 2010; Rosenzweig *et al.*, 2008). Even with estimated probabilities, expert opinions involve substantial subjectivity (Morgan *et al.*, 2009).

Clear guidelines for the presentation and interpretation of impacts research can strengthen confidence in conclusions about the attribution of biological change to climate change. Guidelines must be sufficiently broad to accommodate the wide variety of research approaches and philosophies represented in the literature on climate impacts, including classic Popperian hypothesis testing and Bayesian methods of inference (Platt, 1964; McCarthy, 2007). Marine climate change impact research occurs within diverse disciplines, including biological oceanography, organismal physiology, community ecology and biogeography, joining biological responses at local, regional and even global scales with a range of climate metrics (Garcia *et al.*, 2014). Though each discipline recognizes principles of the scientific method for drawing inference, these principles take various forms when applied to the particular problem of attributing biological change. Comparing the strengths and limitations of inference across this diverse literature is not straightforward, and is further complicated by the wide range of approaches used to communicate the logic of inference. Here, we outline a framework to guide the communication and assessment of inferences in climate change impact studies with the goal of strengthening conclusions about how climate affects biological systems. Using this framework, we develop a specific index, the confidence

index ( $C_{\text{index}}$ ), to assess confidence in conclusions concerning attribution in a sample of marine climate impact studies and to identify areas of strength and those that need improvement in future studies. We analyse current practices, finding that researchers can maximize confidence in conclusions concerning attribution by using prior expectations based on scientific evidence from theory, models, experiments and historical data, along with quantitative analyses of appropriate time-series data.

## A FRAMEWORK TO GUIDE COMMUNICATION OF INFERENCE IN CLIMATE IMPACT STUDIES

Strong inference that climate change has caused a biological response rests on three pillars of the scientific method (Fig. 1). First is a statement of an evidence-based prior expectation for how climate factors affect biological patterns. Second are data: appropriate climate and biological data must be available for testing expectations. Third is quantitative analysis to detect change through time and facilitate the inference of causality. An example consistent with this framework is the conclusion that in the Northern Hemisphere warming associated with climate change has contributed to the reduction of glass eel populations to small fractions of their historical abundances (Bonhommeau *et al.*, 2008). Bonhommeau *et al.* (2008) formulated the expectation that ocean warming has reduced the availability of food for eel larvae and has thus reduced recruitment rates of young eels. Their expectation is based on independent historical evidence of correlated declines in food availability and eel recruitment. An



**Figure 1** A framework outlining the basic scientific process through which a question about attribution of biological change to climate change can be answered, drawing on evidence from theories, experiments and historical data. These evidence-based expectations should guide the choice and use of appropriate data to test the expectation, as well as statistical analyses that allow researchers to distinguish a change through time due to climate from no change, autocorrelation, non-stationarity or response to an alternative driver. Successive iteration of this process reduces the set of plausible expectations and/or builds confidence in a single expected change.

**Table 1** Examples of general and specific prior expectations for the effects of climate change on marine ecological systems based on evidence from theory, experimental and palaeoecological or historical data independent of observed climate change impacts.

Expectation to be tested against time-series data	Theory	Experiment	Palaeo
<b>General expectation 1:</b> Ocean warming shifts species geographic ranges to higher latitudes and deeper water	Cheung <i>et al.</i> (2010)		
(i) Physiological temperature dependence constrains species range		Stillman & Somero (2000), Kuo & Sanford (2009)	Greenstein & Pandolfi (2008)
(ii) Range limits are more dependent on ocean currents than temperature	Gaylord & Gaines (2000)		
(iii) Species geographic ranges change as species interactions at the range borders change	Poloczanska <i>et al.</i> (2008)	Harley (2003)	
<b>General expectation 2:</b> OA reduces the abundance of calcifiers			
(i) OA reduces calcification causing dissolution of CaCO <sub>3</sub> shells in marine calcifiers	Orr <i>et al.</i> (2005)	Gazeau <i>et al.</i> (2007)	Moy <i>et al.</i> (2009)
(ii) OA has a greater effect on calcifiers that produce aragonite and high magnesium–calcite forms of CaCO <sub>3</sub> than those that produce low magnesium–calcite forms of CaCO <sub>3</sub>		Ries <i>et al.</i> (2009)	Kiessling & Simpson (2011)
(iii) Calcifiers that regulate pH levels at the site of calcification were able to maintain calcification rates		de Beer & Larkum (2001), Al-Horani <i>et al.</i> (2003)	Knoll <i>et al.</i> (2007)
<b>General expectation 3:</b> Climate change reduces population connectivity			
(i) Larval development times are shorter in warmer water, reducing potential distance dispersed	O'Connor <i>et al.</i> (2007)	Houde (1989), Pepin (1991)	
(ii) OA reduces larval size for many marine invertebrates, reducing potential survival and larval duration		Kurihara (2008)	
(iii) Warming and OA do not reduce connectivity because ocean currents control dispersal distance	Gaylord & Gaines (2000)		

OA, ocean acidification.

A longer list of expectations derived from the marine climate literature is presented in Appendix S3.

alternative causal pressure – overfishing – was considered and discounted as the only driver of synchronized global declines in eel recruitment. The authors analysed data on sea surface temperature and marine primary production at a spatial (gridded ocean basin) and temporal scale (four decades) sufficient to detect the signal of ocean warming associated with climate change (Henson *et al.*, 2010), and they tested for responses in eel recruitment data at multiple sites within the range of each species. They then used appropriate statistical analyses to detect a temperature-driven signal in eel recruitment throughout their large geographic ranges. Together, these pillars of the framework allowed Bonhommeau *et al.* (2008) to provide strong support for the conclusion that ocean warming, and not confounding factors such as autocorrelation or overfishing, is the primary cause of the globally declining abundance of glass eels.

Climate impacts research should begin with prior expectations, or statements about a biological response to climate change that can be evaluated against data. Expectations take different forms in different philosophies of inference. In a Popperian hypothesis testing approach, an expectation could be a falsifiable hypothesis (Platt, 1964). In Bayesian inference,

expectations take the form of prior distributions (prior knowledge) that are used in combination with observed data to explore a hypothesis (McCarthy, 2007). Predictions are specific deductions from hypotheses or conceptual frameworks. We use the term ‘expectation’ here inclusively, and emphasize that the important step is to formally invoke evidence in the statement of a logical relationship between climate and biological change, but the particular expression of that statement may vary among statistical approaches or scientific disciplines. Importantly, our use of ‘expectation’ is not meant to convey predisposition toward a particular outcome such as a signal of climate change in biological data.

Informative expectations draw upon multiple, independent lines of evidence, and specify a relationship between climate change, the biological response and natural variability in the climate factor and biological variable that might be independent of climate change (Table 1, Appendix 3 in Supporting Information, Fig. 1). Expectations may specify the direction of response (i.e. in Bonhommeau *et al.* (2008) warming will decrease the availability of food for eel larvae) or may more generally specify a climate effect, in contrast to no climate effect (e.g. Hsieh *et al.*

(2005) state that ocean climate change should shift the geographic ranges of fish species). Prior expectations are articulated early in the research process, prior to data analysis and ideally when the study is designed. Explanations of climate impacts often appear as post hoc interpretations without reference to prior expectations, and post hoc interpretations do not provide an equivalent level of support for conclusions (Crisp *et al.*, 2011).

Despite the importance of prior expectations in scientific inference (Platt, 1964; Crisp *et al.*, 2011), surprisingly little attention has been paid to procedures for generating and testing prior expectations in climate change ecology (Rijnsdorp *et al.*, 2009). The first step is the development of prior expectations based on theoretical, empirical or historical evidence (Fig. 1, Table 1). Theoretical evidence takes the form of predictions derived from logical frameworks, sometimes formalized mathematically. Theories can guide the generation of expectations, with the advantage that theoretical frameworks mechanistically link biological change to climate change in a series of logical, deductive steps, allowing each step to be identified, evaluated and tested. For example, metabolic theory relates biological rates to temperature and produces predictions for non-intuitive responses to warming such as stable abundance, despite effects of warming on productivity (O'Connor *et al.*, 2011).

Expectations can be based on experimental evidence, especially when theory is limited, as in the case of ocean acidification (Table 1). Experiments may provide a quantitative estimate of the magnitude and direction of how climatic factors affect biological processes, while controlling for other variables. Experimental manipulations of multiple factors can also inform expectations about synergistic or antagonistic effects (Crain *et al.*, 2008), and may provide insight into the shape of functional responses (Zavaleta *et al.*, 2003). For example, in marine systems, numerous experiments have shown that lower pH negatively affects calcification rates in corals (Ries *et al.*, 2009). These findings suggest that ocean acidification may have reduced the abundance of coral, though the time-series data on pH are not yet long enough to test for an acidification effect over time. When generating expectations for changes through time, experimental evidence should be used with care due to the spatial and temporal constraints of experimental conditions.

Finally, expectations for how modern climate change affects species distributions and relative abundances can be based on historical evidence, including palaeontological and archaeological evidence, encompassing centuries to hundreds of millions of years. For example, an expectation that tropical coral reefs will shift their geographic range poleward to track ocean warming is based on evidence that warmer sea surface temperatures during the last interglacial period (125 ka) resulted in Pleistocene coral reefs in what is now the temperate zone of Western Australia (Table 1; Greenstein & Pandolfi, 2008). Expectations based on historical data do not, however, always provide mechanistic understanding nor do they consider synergies between concurrent anthropogenic pressures that may not have co-occurred in the past. Palaeontological studies have the advantage of long time series where palaeobiological and palaeoclimatic proxy data can be assessed simultaneously, though challenges include

dealing with gaps in data, variability in preservation of different groups of organisms and often resolution of data that differs from the temporal resolution of the biological response to environmental change.

Importantly, evidence for generating expectations must be independent of the data used to test the expectation. For example, Bonhommeau *et al.* (2008) cite historical evidence that eels eat plankton, together with historical evidence that the abundance of plankton has changed with temperature. These datasets are distinct from those used to test for a recent decline in eel recruitment. Similarly, Nye *et al.* (2009) cite historical evidence of shifts in the abundance and distribution of fish species off the Canadian east coast to support a set of directional expectations for how sea surface temperatures should affect fish distributions off the coast of the north-eastern United States. In each case, the expectation for biological change is not based on the recent historical trend in the same system.

Carefully matching the expectation with available data allows inferences to build on the first and third pillars (Fig. 1). For any expected relationship between climate and a biological response, the spatial and temporal scale of the underlying process prescribes the scale and resolution of data appropriate for testing the expectation (Henson *et al.*, 2010). Expectations can also guide the selection of climate metrics, an important step in analysing combined local and regional effects (Garcia *et al.*, 2014). Sufficient data are often not available to properly test an evidence-based expectation. In these cases, the expectation could be restated, perhaps more generally (Table 1), to be testable with the data available. For example, in marine systems, ecological time series are often shorter in duration and coarser in resolution than required to test most expectations concerning ecological impacts of climate change (Henson *et al.*, 2010). Henson *et al.* (2010), a rare example of an impact study that determined data requirements for detecting a climate change response prior to testing, calculated that attribution of observed changes in primary production in oceanic gyres to anthropogenic climate change requires a minimum of 30 years of ocean colour data, whereas the principal source of ocean colour data (a satellite) had only been operating for 14 years. Many marine biological responses have been attributed to climate change with datasets that are too short to deliver conclusive tests. When long-term, continuous datasets are absent, several datasets of shorter duration may be concatenated to create a longer time series (Poloczanska *et al.*, 2008).

Finally, expectations can be used to identify appropriate quantitative tests (Fig. 1). Statistical analyses are necessary to detect a change in particular climate and ecological variables, and to test relationships between them whilst considering confounding factors. Testing expectations against observational data presents many challenges. The simple expectation that climate change is causing biological change implies an expectation that internal variability, other biotic or abiotic factors, or non-stationarity (the tendency of time series to exhibit temporal trends in their statistical properties) are not the only causes of the observed change. Consequently, advanced time-series analyses are often required to detect and attribute the impacts of climate change (Pyper &

Peterman, 1998; Brown *et al.*, 2011). Correlation analysis is commonly used in climate change studies, but this method has rarely accounted for temporal or spatial autocorrelation, non-stationarity or other important time-series issues. Failure to control for these potentially confounding factors introduces problems when assessing the statistical significance of correlations between biological and climate data (Pyper & Peterman, 1998; Brown *et al.*, 2011). Model-based methods of statistical analysis, such as generalized additive modelling and structural equation modelling, can account for multiple drivers of change and autocorrelation simultaneously, allowing researchers to distinguish the contribution of climate change and its interaction with other drivers (Brown *et al.*, 2011). Testing prior expectations that are mechanistically meaningful is important, because neither correlation nor more sophisticated statistical tests alone can determine causation. For instance, if two variables such as fishing pressure and warming are strongly correlated, it will be impossible to distinguish their effects on a species' distribution using statistical methods, whereas it may be possible to exclude one driver from analyses on the basis of a prior expectations for the direction and strength of effect.

## METHODS

### Application of the framework: a quantitative assessment of current practices

The framework outlined here provides guidelines for strengthening inferences drawn from studies attempting to attribute an observed ecological change to climate change (Fig. 1). These guidelines may be applied when designing, reporting or interpreting a climate change impact study, and when synthesizing studies in impact assessments or meta-analyses. To facilitate such applications, we developed a confidence index ( $C_{\text{index}}$ ) that applies a numerical score to individual conclusions about impacts. We used the  $C_{\text{index}}$  to assess current practices in the marine biological literature, and to compare confidence in conclusions about the impact of climate change across taxonomic, geographic and response-type groups. Although application of some components of the framework, such as the minimum number of years required (Table 2, Appendix S1), may be specific to marine climate change, the  $C_{\text{index}}$  could be adapted easily to other systems by changing the specific values of some scoring categories.

### Development of the confidence index

The  $C_{\text{index}}$  measures confidence in scientific inferences derived from climate change impact studies based on scores derived from the scientific method and the spatial and temporal scale of climate change in marine systems (Table 2). Total  $C_{\text{index}}$  scores reflect confidence in the strength of relationships between climate change and biological variables (Fig. 1) based on the type of information typically available for reporting. The  $C_{\text{index}}$  includes an expectation score ( $S_E$ ) that is based on whether a prior expectation was clearly stated for how the biological vari-

able may respond to climate change, whether expectations invoked specific evidence for putative climate pressures and whether expectations were stated for alternative (non-climate-change) causes or confounding factors (Appendix S1). A data score ( $S_D$ ) is based on whether the spatial and temporal extent of the data is well-matched to the scale of relevant climate change (Henson *et al.*, 2010). Scores reflect the number of years in the dataset, the number of sites and the spatial extent of observations (local, regional). A quantitative score ( $S_Q$ ) is based on the statistics used to test (or account for) change through time, autocorrelation and alternative stressors (Brown *et al.*, 2011).

The expectation, data and statistics scores (and their components, Table 2) can be consolidated into scores for detection ( $D = S_D + S_{Q\text{-iii}}$ ) of biological change and understanding ( $U = S_E + S_{Q\text{-iv-v}}$ ) of how the biological change relates to climate change (Table 2) (Kunkel *et al.*, 2013). Detection implies a prior expectation of change in the system, and consequently requires data and statistical procedures that are sufficient to distinguish change from a null pattern and variability inherent in the data, such as natural variation in population size (Brown *et al.*, 2011). Understanding relationships between biological change and climate change (or an alternative causal factor) requires the testing of biological and climate data against prior expectations for how the biological system should respond to observed climate change as well as considering and testing for alternative causal factors. Together, high  $D$  and  $U$  scores convey high confidence in conclusions concerning attribution, and thus may indicate studies that are most relevant to synthetic impacts assessments.

The maximum possible  $C_{\text{index}}$  score, 18, describes a data-rich statistical test of a clearly stated, evidence-based expectation for how climate change affects a biological response, including a concurrent test for the role of alternative causal factors. Notably, confidence in conclusions concerning the cause of a biological change is independent of a study's findings. A conclusion may earn high confidence regardless of whether a biological response is attributed to climate change, to other causal factors, or no change was detected. For example, scores do not indicate whether the prior expectation was supported by the statistical test: a high-scoring study may find that climate change is not driving biological change, and the methods used may convey high confidence in this conclusion.

### Analysis of $C_{\text{index}}$ scores

We applied the  $C_{\text{index}}$  to 1735 biological time series, or 'observations', in 208 marine climate change impact studies to evaluate overall confidence in assessments of marine biological responses to climate change (Appendix S2). Here, we consider an 'observation' to be a time series of a single biological variable (i.e. abundance or phenology).

For this dataset, Poloczanska *et al.* (2013) determined whether each observation was consistent with the original authors' stated or implied expected impact of climate change. The characterization of 'consistent', 'not consistent' or 'no change' is not included in the  $C_{\text{index}}$ , which remains neutral on

**Table 2** The confidence index ( $C_{\text{index}}$ ) is composed of three subscores: (a) the statement and testing of prior expectations for climate impacts as well as for alternative causal factors, (b) the data available to test the expectations, and (c) the quantitative or statistical analyses applied. The  $C_{\text{index}}$  score can be objectively assessed in a climate impact study and is calculated for a single observed change (e.g. a biological time-series dataset) and the associated analyses and inferences surrounding the cause of that change. The two columns indicate scores that contribute to detection (D) of change and understanding (U) of how climate affects change, respectively (Kunkel *et al.*, 2013).

Detection (D)	Understanding (U)	Score
<b>(a) Expectations (<math>S_E</math>)</b>		Max. = 6
	(i) 0 = no clear prior expectation stated +2 = prior expectation is clearly stated	
	(ii) +2 = study invokes specific evidence to support prior expectation	
	(iii) +1 = prior expectations are based on multiple lines of evidence	
	(iv) +1 = prior expectations for alternative causal factors are articulated or confounding factors are discounted	
<b>(b) Data (<math>S_D</math>)</b>		Max. = 6
Temporal		
	(i) 0 = $\leq 10$ years of data +1 = 11–25 years of data +2 = $> 25$ years of data + 1 = data spans $> 30$ years	
	(ii) +1 = continuous data (annual)	
Spatial		
	(iii) 0 = single site study +1 = multiple site study	
	(iv) 0 = local ( $\leq 1000 \text{ km}^2$ ) +1 = regional ( $> 1000 \text{ km}^2$ ) +1 = broad scale ( $> 100,000 \text{ km}^2$ )	
<b>(c) Quantitative analysis (<math>S_Q</math>)</b>		Max. = 6
(i) +1 = biological trend was statistically tested to distinguish from no trend	(iv) +2 = relationship between biological variable and climatic variable was statistically tested	
(ii) +1 = a trend in a climatic variable was statistically tested to distinguish from no trend	(v) +1 = alternative causal factors tested	
(iii) +1 = temporal/spatial autocorrelations have been considered		
Max. = 9	Max. = 9	Max. = 18

the outcome of an analysis with regard to the direction or magnitude of its concluded biological response to climate change. After calculating  $C_{\text{index}}$  scores for each observation, we matched these scores with the previous classification of consistency.

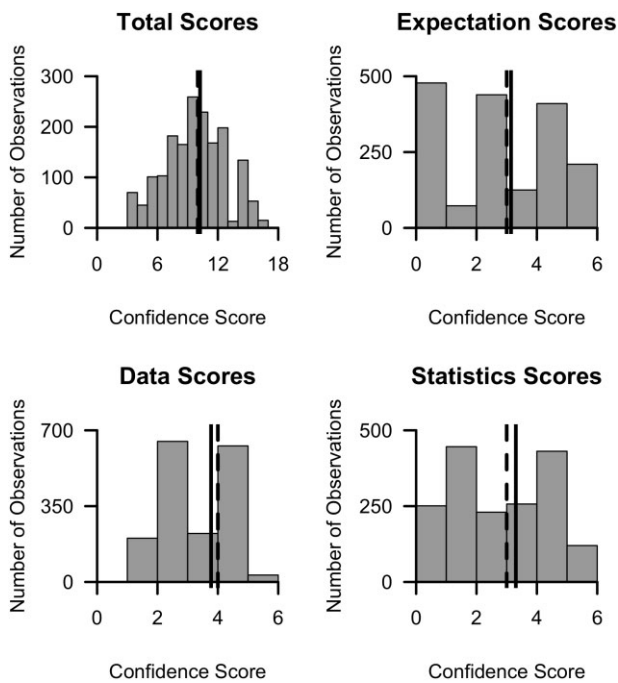
$C_{\text{index}}$  total scores were not normally distributed (Shapiro–Wilks test:  $W = 0.98$ ,  $P < 0.001$ ), and we tested for differences in scores among ocean, taxonomic groups and response-type groups using Kruskal–Wallis nonparametric tests. Although the grouping of observations within studies represents a kind of pseudoreplication, we chose not to model the study-level variation. Our objective was to assess confidence in conclusions at the level of observation because these are typically the units used in subsequent meta-analyses and syntheses that inform a collec-

tive understanding of climate impacts, and study-level factors are likely to contribute to these patterns.

## RESULTS

### Results of $C_{\text{index}}$ analysis in marine biological studies

We found a wide range of  $C_{\text{index}}$  scores (from 1 to 16) across our set of biological observations (Fig. 2). Most observations scored highly in only one category (e.g. data,  $S_D$ ), and no observation scored highly in all three categories. The maximum possible scores for expectations ( $S_E$ ), data ( $S_D$ ) and statistics ( $S_Q$ ) were achieved by 210 (12%), 32 (2%) and 120 (7%) observations, respectively (Fig. 2). More than 35% of observations were not



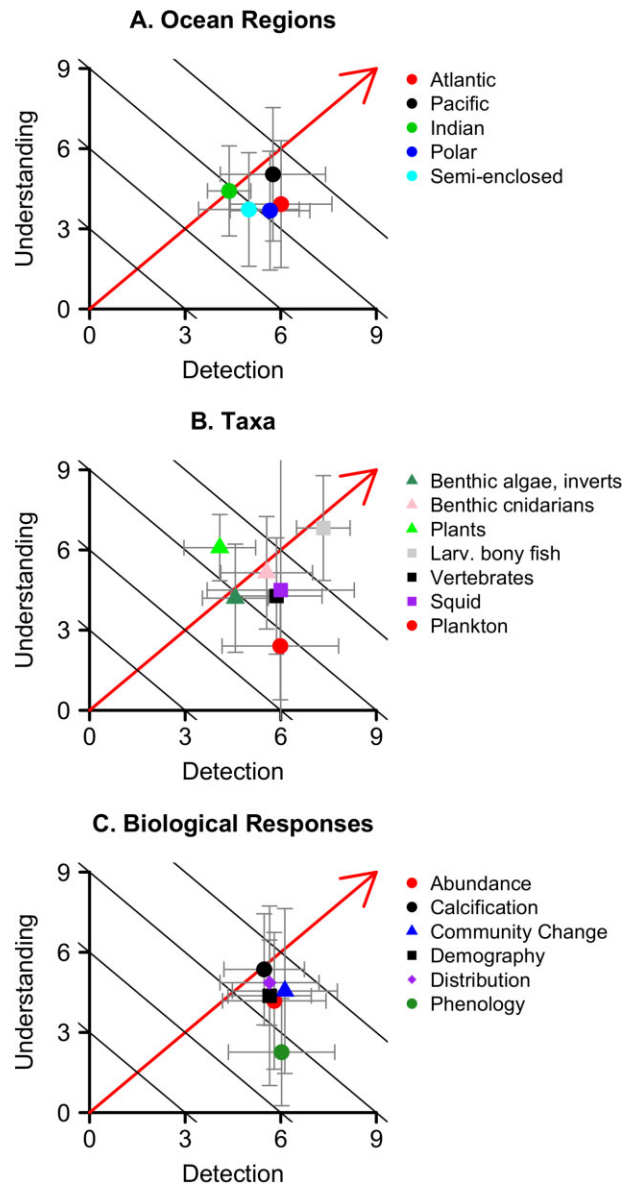
**Figure 2** Distribution of  $C_{\text{index}}$  scores for marine climate impact studies, along with distributions of subscores for prior expectations, data and quantitative scores. A solid black line indicates the mean score and a dashed line indicates the median score.

reported in association with a prior expectation, and only 35% of statistical tests involved a test of an alternative driver of the observed result.

The highest scoring observations (330 observations, or 19%, across 37 studies) earned scores of six or more out of a possible nine in both detection ( $D$ ) and understanding ( $U$ ) (Fig. 3a, Appendix 2). Of these, 91% articulated a prior expectation and 86% articulated an expectation for an alternative causal factor; 100% statistically tested a change in climate but only 63% statistically tested for a change in an alternative causal factor; 90% used time series with more than 30 years of data; and 98% were multisite comparisons. Biological change was considered by the original authors to be consistent with expected effects of climate change in 226 (68%) of these high-scoring observations.

Biological change through time was not detected for all observations (Poloczanska *et al.*, 2013) (Table 3). Observations for which change was detected (consistent with climate change) earned the lowest average  $C_{\text{index}}$  scores (Kruskal–Wallis test:  $\chi^2 = 23.2$ , d.f. = 2,  $P < 0.001$ ; Table 3) due to lower expectation and statistics scores (Table 3). When multiple species responses were reported in the same study, observations tended to earn higher scores than observations reported in studies that only considered one or two species (Kruskal–Wallis test:  $\chi^2 = 8.93$ , d.f. = 1,  $P < 0.003$ ). The difference is due to statistics scores (Kruskal–Wallis test:  $\chi^2 = 11.54$ , d.f. = 1,  $P < 0.0007$ ) and data scores (Kruskal–Wallis test:  $\chi^2 = 10.96$ , d.f. = 1,  $P < 0.001$ ).

To assess patterns in confidence of inferred climate change responses, we considered only observations for biological



**Figure 3** The confidence index ( $C_{\text{index}}$ ) quantifies confidence in attribution of an observed biological change to climate change. Attribution requires formal detection of change in a biological system, as well as formal understanding of how climate relates to the observed change (Table 1). Increasing total confidence is indicated by the red arrow. Black lines represent isoclines. Higher scores convey higher confidence in the conclusions of a study.  $C_{\text{index}}$  scores vary with available data and research approaches, and influence confidence in the current scientific attribution of biological responses to climate change as measured by the mean ( $\pm$  SD)  $C_{\text{index}}$  score for studies grouped by ocean region (a), taxon (b) and biological response (c).

change consistent with climate change, as determined by a study's authors ( $n = 1098$  observations; Appendix 2). Scores differed among ocean regions (Kruskal–Wallis test:  $\chi^2 = 29.23$ , d.f. = 4,  $P < 0.001$ , Table 4), with the Pacific and Atlantic oceans tending to earn higher confidence scores than other regions

**Table 3** Mean confidence index ( $C_{\text{index}}$ ) scores and results of Kruskal–Wallis test for differences in scores among groups of observations considered by a study's original authors to be consistent or opposite to expected, or no change. Letter superscripts indicate significant differences among groups, based on Mann–Whitney pairwise comparisons of group means and variances. Bold values indicate significant  $P$ -values relative to criteria ( $\alpha$ -levels indicated in the table) reflecting Bonferroni adjustments for multiple tests.

$i$	$j$	Difference in means ( $i - j$ )	Difference in median ( $i - j$ )	$W$	$P$
$C_{\text{index}}$ score					
Consistent	No change	9.97–10.58 = –0.62	10–11 = –1	196,388	$\alpha = 0.02$
	Opposite to expected	9.97–10.88 = –0.91	10–11 = –1	105,389.5	<b>&lt; 0.01</b>
No change	Opposite to expected	10.58–10.88 = –0.29	11–11 = 0	44,775.5	0.48
Expectation score					
Consistent	No change	2.97–3.30 = –0.33	3–3 = 0	207,683.5	$\alpha = 0.02$
	Opposite to expected	2.97–3.69 = –0.72	3–5 = –2	97,831	<b>&lt; 0.01</b>
No change	Opposite to expected	3.30–3.69 = –0.39	3–5 = –2	38,808.5	<b>&lt; 0.01</b>
Data score					
Consistent	No change	3.76–3.76 = 0.01	4–3 = 1	224,198	$\alpha = 0.02$
	Opposite to expected	3.76–3.89 = –0.13	4–4 = 0	116,087	0.78
No change	Opposite to expected	3.76–3.89 = –0.12	3–4 = –1	43,901.5	0.14
Stats score					
Consistent	No change	3.22–3.52 = –0.30	3–3 = 0	202,323.5	$\alpha = 0.02$
	Opposite to expected	3.22–3.29 = –0.07	3–3 = 0	120,108.5	<b>&lt; 0.02</b>
No change	Opposite to expected	3.52–3.29 = 0.23	3–3 = 0	50,626.5	0.50

(Fig. 3, Table 4). Taxonomic groups also differed (Kruskal–Wallis test:  $\chi^2 = 233.03$ , d.f. = 6,  $P < 0.001$ ), and the highest confidence scores were observed for larval bony fish, followed by vertebrates and benthic cnidarians (Fig. 3c, Table 4). Phenological responses earned significantly lower  $C_{\text{index}}$  scores than abundance, distribution, demography, calcification and community change (Fig. 3, Table 4) (Kruskal–Wallis test:  $\chi^2 = 53.89$ , d.f. = 5,  $P < 0.001$ ).

Detection scores, primarily reflecting the availability of data, explain differences in confidence scores among ocean regions (Fig. 3a). In contrast, understanding scores for how climate influences different responses explain differences among biological response types (Fig. 3c): phenological observations received detection scores similar to those of other observation types, but earned lower understanding scores (Fig. 3d). In general, confidence scores for observations of benthic organisms were limited by detection scores, while confidence scores for observations of planktonic organisms scored highly for detection but weakly for understanding (Fig. 3b).

## DISCUSSION

### Recommendations for increasing confidence in climate impact studies

As quantified by the  $C_{\text{index}}$ , observations of marine biological change are attributed to climate change with, on average, moderate confidence. Our findings suggest high confidence in attribution to climate change for changes involving larval bony

fish, Pacific and Atlantic ocean regions and distributional, calcification and community-level responses. These findings differ from a confidence ranking based only on sampling effort (number of observations) (Table 4), and suggest that estimates of confidence based on vote counting do not necessarily align with estimates based on the integrated use of data, statistics and evidence-based expectations.

Several of our findings could reflect subtle publication biases in the impacts literature. When no change was detected in a biological time series, or when observations were contrary to an expected result of climate change, these conclusions earned higher confidence scores than analyses for which change was consistent with expectations. In addition, observations from multispecies assemblages tended to earn higher confidence. Multispecies studies tend to report changes (or no change) through time for all species in a dataset, regardless of consistency with expected impacts of climate change (e.g. Hsieh *et al.*, 2005). These patterns in  $C_{\text{index}}$  scores are consistent with literature bias against studies that show no effect of climate change. Solutions to publication bias include enhancing opportunities for publishing results that are not found to be consistent with climate change (Parmesan & Yohe, 2003) and for individual researchers to determine data needs based on prior expectations and then to report all results, including any that are indeterminate or counter to expectations.

Most observations did not earn high confidence scores. In an effort to highlight priorities for future research, we explore where points were lost and suggest possible solutions to three common problems we observed in our synthesis.



**Table 4** Mean  $C_{\text{index}}$  scores and results of linear models testing for differences in scores for observations of different ocean regions, taxa or biological responses. All observations included here detected change that was considered by a study's original authors to be consistent with climate change. Bold values indicate significant  $P$ -values relative to criteria ( $\alpha$ ) reflecting Bonferroni adjustments for multiple tests, as indicated.

Factor ( <i>i</i> )	<i>n</i>	Mean score	Median score	Contrast ( <i>j</i> )	<i>W</i>	<i>P</i>
Ocean region						$\alpha = 0.005$
Atlantic Ocean	595	9.93	10	Pacific Ocean	73,079.5	0.014
				Polar seas	37,483	0.140
				Semi-enclosed seas	21,362	<b>&lt; 0.003</b>
				Indian Ocean	20,083.5	<b>0.005</b>
Pacific Ocean	274	10.79	10	Indian Ocean	4870.5	<b>&lt; 0.001</b>
				Polar seas	18,805	<b>0.004</b>
				Semi-enclosed seas	10,380	<b>0.002</b>
Polar Seas	116	9.34	10.5	Semi-enclosed seas	3835	0.128
				Indian Ocean	2717.5	0.107
Semi-enclosed Seas	58	8.72	9	Indian Ocean	1684.5	0.604
Indian Ocean	55	8.80	8			
Taxa						$\alpha = 0.002$
Vertebrates	543	10.47	10	Plankton	42,156.5	<b>&lt; 0.001</b>
				Benthic invertebrates	33,362.5	<b>&lt; 0.001</b>
				Larval bony fish	47,790.5	<b>&lt; 0.001</b>
				Benthic cnidarians	10,649	0.130
				Plants	6325	0.920
				Squid	1091	0.990
Plankton	219	8.35	9	Benthic invertebrates	20,556.5	0.176
				Larval bony fish	20,526	<b>&lt; 0.001</b>
				Benthic cnidarians	5481	<b>&lt; 0.001</b>
				Plants	1809.5	0.026
				Squid	391	0.714
Benthic invertebrates	174	8.77	9	Larval bony fish	1323.5	<b>&lt; 0.001</b>
				Benthic cnidarians	1337	<b>&lt; 0.001</b>
				Plants	1414	0.021
				Squid	340	0.941
Larval bony fish	101	14.44	15	Benthic cnidarians	350.5	<b>&lt; 0.001</b>
				Plants	2128.5	<b>&lt; 0.001</b>
				Squid	199	0.965
Benthic cnidarians	34	10.75	11	Plants	497.5	0.075
				Squid	67	0.980
Plants	23	9.82	9	Squid	46	1
Squid	4	10.50	10.5			
Observation type						$\alpha = 0.003$
Abundance	487	9.98	10	Distribution	77,951.5	<b>0.003</b>
				Phenology	45,683	<b>&lt; 0.001</b>
				Demography	10,307	0.865
				Community change	6919.5	0.434
				Calcification	4817.5	0.077
Distribution	363	10.51	11	Phenology	38,386.5	<b>&lt; 0.001</b>
				Demography	1752.5	0.366
				Community change	5535	0.880
				Calcification	4830.5	0.586
Phenology	149	8.29	9	Demography	4107.5	0.005
				Community change	3010	0.007
				Calcification	2838.5	<b>&lt; 0.001</b>
Demography	43	10.02	10	Community change	715.5	0.593
				Calcification	6160	0.355
Community change	31	10.68	9	Calcification	426.5	0.522
Calcification	25	10.84	11			

*Data may be insufficient to test an expectation of climate-driven change*

Weak  $S_D$  scores were common (Fig. 2). Studies from the Indian Ocean or focusing on benthic taxa, for example, scored low on detection (Fig. 3) because relatively few long-term datasets exist or have been analysed in the context of the impacts of climate change. One solution is to increase investment in time-series datasets so that these may be available for future analysis (Reichman *et al.*, 2011). In addition, efforts to support public archiving of data will ensure that existing data can be used well into the future (Wolkovich *et al.*, 2012; Vines *et al.*, 2014).

*Statistical analysis may have been inappropriate or had insufficient power to detect change and test for climate-driven causes*

Fewer than 10% of analyses of time series scored the maximum possible statistics scores; correlation analyses were most commonly used (Brown *et al.*, 2011). Brown *et al.* (2011) reviewed an expanded version of the database we have analysed here and provided recommendations on the most robust approaches for tackling statistical challenges. More complex methods of detecting causation such as the convergent cross mapping (CCM) approach can help to identify causation by the increased ability of independent variables to predict dependent variables – and vice versa in the case of non-external independent variables – over time (Sugihara *et al.*, 2012). This lack of convergence can cause re-examination of time series to produce conflicting results (Myers, 1998). CCM proves useful in the case of external forcing on two non-coupled variables, as cross-mapping will show no evidence of convergence between the two biological variables thus demonstrating no causation between them, although they may appear correlated (Sugihara *et al.*, 2012).

*Prior expectations for climate impacts and alternative explanations may be lacking, vague or insufficient in the final manuscript*

Many studies (119 of 208 in our sample) lacked a clearly stated expectation for how climate change may affect the study system. Possible explanations for not stating a prior expectation in a climate impact study include the assumption by the authors (or reviewers) that readers possess expert knowledge sufficient to independently assume an expectation, convention in certain disciplines or cultures to de-emphasize prior expectations in scientific reporting and analyses, and uncertainty in the expected biological response to climate change.

When data are limited, the statement of prior expectations can strengthen confidence in attribution. For one iconic taxon, reef-building corals, for which extensive time-series data are generally lacking, climate impacts are attributed with moderate confidence in our database (Fig. 3). Studies of benthic cnidarians, including reef-building corals, scored higher  $C_{index}$  values than other taxa with moderate detection scores (Fig. 3b),

reflecting the use of evidence-based expectations to support conclusions about how climate has affected corals (De'ath *et al.*, 2009; Pandolfi *et al.*, 2011). Furthermore, prior expectations allow for proper selection of climate metrics when the vulnerability of a species to climate change is species or system specific (Garcia *et al.*, 2014).

Our analysis quantified confidence in attribution within primary research studies, and we used these results to identify strengths and opportunities for improvement in climate impact studies. The framework we have developed, and the associated confidence index, does not suggest that a greater number of observations confers greater confidence in the conclusions. The  $C_{index}$  and the framework from which it is derived suggest that greater confidence is achieved by integrating independent evidence for climate impacts into expectations that are then statistically tested against time-series data. Other approaches for assessment and attribution, such as meta-analyses and hierarchical models, pooled observations and consolidate shared information across groups to reduce uncertainty and thereby strengthen inference (Hedges *et al.*, 1999; Gelman, 2006). These statistical approaches are necessary to determine the strength of evidence provided by a set of observations. The  $C_{index}$  can be used within synthetic statistical frameworks as a weighting factor, to give greater influence to observations with a higher confidence score, strengthening cross-study comparisons such as those made by the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports and advancing the study of climate change ecology. Currently the IPCC uses 'expert opinion' linked to a probability rating to determine strength of attribution to climate change within meta-analyses. Ranking studies or time series using a numerical scale as we suggest here would reduce the subjectivity that may be inherently present in expert judgements (Morgan *et al.*, 2009) such that the confidence in attribution to climate change could be averaged per species, per region or per response type as we have done here. Because observations are ranked in an ordinal manner, greater emphasis can be placed on those observations with greater confidence when defining percentage consistency with climate change, as done in many meta-analyses of climate impacts (e.g. Parmesan & Yohe, 2003; Poloczanska *et al.*, 2013), thus better reflecting the current state of knowledge of the impacts of climate change on biological changes.

## CONCLUSIONS

We have offered the  $C_{index}$  as a more transparent approach to assessing confidence in climate impacts studies, given the variability in presentation and methods in this literature. Strengths of the  $C_{index}$  include its synthesis of evidence from a broad pool of scientific information to formulate testable expectations consistent with the scientific method, and some degree of transparency in how confidence is assessed and compared. The approach therefore overcomes some of the limitations associated with less transparent methods based on subjective expert opinions.

Nonetheless, reproducibility remains a major challenge in climate impacts research and assessments. The challenge of

reproducibility is two-fold. First, given the same data, are conclusions and confidence assessments reproducible? Second, are findings reproducible when longer or more extensive datasets become available (Myers, 1998)? Our confidence index is one attempt to develop a confidence assessment procedure that may be more reproducible than expert opinions (Morgan *et al.*, 2009) due to the clearly defined categories and transparent assignment of scores. The  $C_{\text{index}}$  only indirectly addresses a second challenge of reproducibility, by encouraging communication of logic such that research processes and conclusions might be reproducible even with more data. However, if historical data and analytical code are not made available with publications, a deeper understanding of biological change through time may be hampered, despite the availability of more data (Wolkovich *et al.*, 2012; Vines *et al.*, 2014).

We found that many studies present compelling, but not conclusive, evidence that climate change more than other factors has affected marine biological systems. Confidently attributing biological change to climate change is essential to the application of scientific research to decision making. To improve this process, we have outlined an approach for assessing confidence in attribution at the level of the research analysis. Future investment in long-term datasets and the basic science to support evidence-based prior expectations for climate impacts will directly strengthen confidence in impacts studies and assessments.

## ACKNOWLEDGEMENTS

We are grateful to Dr Sylvain Bonhommeau and the anonymous referees for thoughtful comments and suggestions that improved this work. We acknowledge financial support from the National Center for Ecological Analysis and Synthesis (NCEAS) to A.J.R. and E.S.P.

## REFERENCES

- Al-Horani, F., Al-Moghrabi, S. & de Beer, D. (2003) The mechanism of calcification and its relation to photosynthesis and respiration in the scleractinian coral *Galaxea fascicularis*. *Marine Biology*, **142**, 419–426.
- Bonhommeau, S., Chassot, E., Planque, B., Rivot, E., Knap, A.H. & Le Pape, O. (2008) Impact of climate on eel populations of the Northern Hemisphere. *Marine Ecology Progress Series*, **373**, 71–80.
- Brown, C.J., Schoeman, D.S., Sydeman, W.J., Brander, K., Buckley, L.B., Burrows, M., Duarte, C.M., Moore, P.J., Pandolfi, J.M., Poloczanska, E., Venables, W. & Richardson, A.J. (2011) Quantitative approaches in climate change ecology. *Global Change Biology*, **17**, 12, 3697–3713.
- Cheung, W.W.L., Lam, V.W.Y., Sarmiento, J.L., Kearney, K., Watson, R., Zeller, D. & Pauly, D. (2010) Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, **16**, 24–35.
- Crain, C.M., Kroeker, K. & Halpern, B.S. (2008) Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters*, **11**, 1304–1315.
- Crisp, M.D., Trewick, S.A. & Cook, L.G. (2011) Hypothesis testing in biogeography. *Trends in Ecology and Evolution*, **26**, 66–72.
- De Beer, D. & Larkum, A. (2001) Photosynthesis and calcification in the calcifying algae *Halimeda discoidea* studied with microsensors. *Plant, Cell and Environment*, **24**, 1209–1217.
- De'ath, G., Lough, J.M. & Fabricius, K.E. (2009) Declining coral calcification on the Great Barrier Reef. *Science*, **323**, 116–119.
- Garcia, R.A., Cabeza, M., Rabhek, C. & Araujo, M.B. (2014) Multiple dimensions of climate change and their implications for biodiversity. *Science*, **344**, 486–497.
- Gaylord, B. & Gaines, S.D. (2000) Temperature or transport: species ranges mediated solely by flow. *The American Naturalist*, **155**, 769–789.
- Gazeau, F., Quiblier, C., Jansen, J., Gattuso, J.P., Middleburg, J. & Heip, C. (2007) Impact of elevated CO<sub>2</sub> on shellfish calcification. *Geophysical Research Letters*, **34**, L07603.
- Gelman, A. (2006) Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics*, **48**, 432–435.
- Greenstein, B.J. & Pandolfi, J.M. (2008) Escaping the heat: range shifts of reef coral taxa in coastal Western Australia. *Global Change Biology*, **14**, 513–528.
- Harley, C.D.G. (2003) Abiotic stress and herbivory interact to set range limits across a two-dimensional stress gradient. *Ecology*, **84**, 1477–1488.
- Hedges, L.V., Gurevitch, J. & Curtis, P.S. (1999) The meta-analysis of response ratios in experimental ecology. *Ecology*, **80**, 1150–1156.
- Hegerl, G.C., Zwiers, F.W., Braconnot, P., Gillett, N.P., Luo, Y., Marengo Orsini, J.A., Nicholls, N., Penner, J.E. & Stott, P.A. (2007) Understanding and attributing climate change. *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (ed. by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller), Cambridge University Press, Cambridge, UK.
- Hegerl, G.C., Hoegh-Guldberg, O., Casassa, G., Hoerling, M.P., Kovats, R.S., Parmesan, C., Pierce, D.W. & Stott, P.A. (2010) Good practice guidance paper on detection and attribution related to anthropogenic climate change. *Meeting report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change* (ed. by T.F. Stocker, C.B. Field, D. Qin, V. Barros, G.-K. Plattner, M. Tignor, P.M. Midgley and K.L. Ebi), IPCC Working Group I Technical Support Unit, University of Bern, Bern.
- Henson, S.A., Sarmiento, J.L., Dunne, J.P., Bopp, L., Lima, I., Doney, S.C., John, J. & Beaulieu, C. (2010) Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, **7**, 621–640.

- Houde, E.D. (1989) Comparative growth, mortality and energetics of marine fish larvae: temperature and implied latitudinal effects. *Fishery Bulletin*, **87**, 471–495.
- Hsieh, C., Reiss, C., Watson, W., Allen, M.J., Hunter, J.R., Lea, R.N., Rosenblatt, R.H., Smith, P.E. & Sugihara, G. (2005) A comparison of long-term trends and variability in the populations of larvae of exploited and unexploited fishes in the southern California region: a community approach. *Progress in Oceanography*, **67**, 160–185.
- Kiessling, W. & Simpson, C. (2011) On the potential for ocean acidification to be a general cause of ancient reef crises. *Global Change Biology*, **17**, 56–67.
- Knoll, A., Barnbach, R., Payne, J., Pruss, S. & Fischer, W. (2007) Paleophysiology and end-Permian mass extinction. *Earth and Planetary Science Letters*, **256**, 295–313.
- Kunkel, K.E., Karl, T.R., Brooks, H. *et al.* (2013) Monitoring and understanding trends in extreme storms: state of knowledge. *Bulletin of the American Meteorological Society*, **94**, 499–514.
- Kuo, E.S.L. & Sanford, E. (2009) Geographic variation in the upper thermal limits of an intertidal snail: implications for climate envelope models. *Marine Ecology Progress Series*, **388**, 137–146.
- Kurihara, H. (2008) Effects of CO<sub>2</sub>-driven ocean acidification on the early developmental stages of invertebrates. *Marine Ecology Progress Series*, **373**, 275–284.
- McCarthy, M.A. (2007) *Bayesian methods for ecology*. Cambridge University Press, Cambridge.
- Morgan, M., Dowlatabadi, M., Henrion, M., Keith, D., Lempert, R., McBride, S., Small, M. & Wilbanks, T. (2009) Best practice approaches for characterizing, communicating and incorporating scientific uncertainty in climate decision making. A report by the Climate Change Science Program and the Subcommittee on Global Change Research. National Oceanic and Atmospheric Administration, Washington, DC.
- Moy, A.D., Howard, W.R., Bray, S.G. & Trull, T.W. (2009) Reduced calcification in modern Southern Ocean planktonic foraminifera. *Nature Geoscience*, **2**, 276–280.
- Myers, R.M. (1998) When do environment–recruitment correlations work? *Review in Fish Biology and Fisheries*, **8**, 285–305.
- Nye, J.A., Link, J.S., Hare, J.A. & Overholtz, W.J. (2009) Changing spatial distribution of fish stocks in relation to climate and population size on the northeast United States continental shelf. *Marine Ecology Progress Series*, **393**, 111–129.
- O'Connor, M.I., Bruno, J.F., Gaines, S.D., Lester, S.E., Halpern, B.S., Kinlan, B.P. & Weiss, J.M. (2007) Temperature control of larval dispersal and implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences USA*, **104**, 1266–1271.
- O'Connor, M.I., Gilbert, B. & Brown, C.J. (2011) Theoretical predictions for how temperature affects the dynamics of interacting herbivores and plants. *The American Naturalist*, **178**, 5, 626–638.
- Orr, J.C., Fabry, V.J., Aumont, O. *et al.* (2005) Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, **437**, 681–686.
- Pandolfi, J.M., Connolly, S.R., Marshall, D.J. & Cohen, A.L. (2011) Projecting coral reef futures under global warming and ocean acidification. *Science*, **333**, 418–422.
- Parnesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution and Systematics*, **37**, 637–639.
- Parnesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.
- Pepin, P. (1991) Effect of temperature and size on development, mortality and survival rates of the pelagic early life history stages of marine fish. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 503–518.
- Platt, J.R. (1964) Strong inference: certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, **146**, 347–353.
- Poloczanska, E.S., Hawkins, S.J., Southward, A.J. & Burrows, M.T. (2008) Modeling the response of populations of competing species to climate change. *Ecology*, **89**, 3138–3149.
- Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K.M., Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., O'Connor, M.I., Pandolfi, J.M., Parnesan, C., Schwing, F.B., Thompson, S.A. & Richardson, A.J. (2013) Global imprint of climate change on marine life. *Nature Climate Change*, **3**, 919–925.
- Pyper, B.J. & Peterman, R.M. (1998) Comparison of methods to account for autocorrelation in correlation analyses of fish data. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 2127–2140.
- Reichman, O.J., Jones, M.B. & Schildhauer, M.P. (2011) Challenges and opportunities of open data in ecology. *Science*, **331**, 703–705.
- Ries, J., Cohen, A. & McCorkle, D. (2009) Marine calcifiers exhibit mixed responses to CO<sub>2</sub>-induced ocean acidification. *Geology*, **37**, 1131–1134.
- Rijnsdorp, A.D., Peck, M.A., Engelhard, G.H., Mollmann, C. & Pinnegar, J.K. (2009) Resolving the effect of climate change on fish populations. *ICES Journal of Marine Science*, **66**, 1570–1583.
- Rosenzweig, C., Karoly, D., Vicarelli, M., Neofotis, P., Wu, Q., Casassa, G., Menzel, A., Root, T.L., Estrella, N., Seguin, B., Tryjanowski, P., Liu, C., Rawlins, S. & Imeson, A. (2008) Attributing physical and biological impacts to anthropogenic climate change. *Nature*, **453**, 353–357.
- Stillman, J.H. & Somero, G.N. (2000) A comparative analysis of the upper thermal tolerance limits of eastern Pacific porcelain crabs, genus *Petrolisthes*: influences of latitude, vertical zonation, acclimation, and phylogeny. *Physiological and Biochemical Zoology*, **73**, 200–208.
- Sugihara, G., May, R.M., Ye, H., Deyle, E.R., Fogarty, M. & Munch, S.B. (2012) Detecting causality in complex ecosystems. *Science*, **338**, 496–500.
- Vines, T.H., Albert, A.Y.K., Andrew, R.L., Debarre, F., Bock, D.G., Franklin, M.T., Gilbert, K.J., Moore, J.-S., Renaut, S. &

Rennison, D. (2014) The availability of research data declines rapidly with article age. *Current Biology*, **24**, 94–97.

Wolkovich, E.M., Regetz, J. & O'Connor, M.I. (2012) Advances in global change research require open science at the individual-level. *Global Change Biology*, **18**, 2102–2110.

Zavaleta, E.S., Shaw, M.R., Chiariello, N.R., Mooney, H.A. & Field, C.B. (2003) Additive effects of simulated climate changes, elevated CO<sub>2</sub>, and nitrogen deposition on grassland diversity. *Proceedings of the National Academy of Sciences USA*, **100**, 7650–7654.

Additional references to the data sources may be found in Appendices S1 & S2. [weblink]

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

**Appendix S1** Supplementary methods.

**Appendix S2** Table of confidence index scores for papers included in the analysis.

**Appendix S3** Table of example expectations for different marine systems' response to climate change.

## BIOSKETCHES

**Mary O'Connor** is a marine ecologist researching the drivers of variation in community structure and function, with a particular focus on the role of temperature-dependent metabolism in determining community- and ecosystem-level processes.

**Johnna Holding** studies the metabolic and ecological consequences of changing ocean conditions for marine life.

**Carrie Kappel's** research focuses on quantifying the ways humans depend upon and affect marine species and ecosystems and developing tools and approaches to enhance ocean management.

All authors are climate change ecologists, and have collaborated on various projects related to climate impact arising from the working group Marine Biological Impacts of Climate Change, sponsored by the National Centre for Ecological Analysis and Synthesis, and led by **Anthony Richardson** and **Elvira Poloczanska**.

Editor: José Alexandre Diniz-Filho