Journal of Experimental Marine Biology and Ecology 480 (2016) 1–7

Contents lists available at ScienceDirect

Journal of Experimental Marine Biology and Ecology

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

Predicting the diet of coastal fishes at a continental scale based on taxonomy and body size

German A. Soler [a,*], Graham J. Edgar [a], Rick D. Stuart-Smith [a], Anthony D.M. Smith [b], Russell J. Thomson [a]

[a] Institute for Marine and Antarctic Studies, University of Tasmania, Private Bag 49, Hobart, TAS 7001, Australia
[b] CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, TAS 7001, Australia

A R T I C L E I N F O

Article history:
Received 2 December 2015
Received in revised form 26 March 2016
Accepted 26 March 2016
Available online 6 April 2016

Keywords:
Temperate Australia
Trophic model
Community consumption
Mean prey size
Seagrass

A B S T R A C T

Predicting diet of animals in ecological communities is necessary for a better understanding of trophic links and piecing together food webs to inform ecosystem-based management. A dietary model, *Consume*, was recently developed to predict detailed dietary information for fishes on the basis of fish identity and size. This model was field-tested over a continental scale, predicting community-level consumption for other temperate marine fish communities that differed in species composition and size structure. Using local stomach contents data to field-test predictions, accurate performance of the model was found across 14 locations around southern Australia. Prey type and mean prey size were predicted for fishes at new locations with high accuracy (mean percentage overlap between predicted and actual prey types = 73%; r^2 between predicted and observed mean prey size = 89%) when trained with stomach contents data from subsets of sampled fishes at all locations. Model accuracy dropped, but was still respectable, when using training data only from one location (prey type accuracy = 67%; mean prey size r^2 = 56%). Prey type was more accurately predicted on the basis of consumer body size than species identity, while consumer family identity and size were needed for accurate prediction of mean prey size. The most important factors were evaluated by leaving out predictors (species, genus and family identity; size of consumer; habitat, location, ecoregion and biogeographic province). Exclusion of geographical location information resulted in little loss in accuracy. Our results highlight the need for consideration of consumer body size in trophic models, rather than binning species into functional groups solely on the basis of taxonomy. Application of *Consume* to situations where no dietary information exists, but at least fish family identity and size structure are known, will provide a novel mechanism for testing important ecological hypotheses and assessing trophic consequences of anthropogenically-induced changes in community structure.

1. Introduction

Predation is an important process determining the structure of marine communities (Russ 1980), with critical information on the nature and magnitude of this captured in the diet of predators in the system (Edgar and Shaw 1995a, 1995b). Fish represent major predators in shallow marine systems, and the ability to predict the diet of fishes at a given location should allow improved understanding of the ecological dynamics of the community (Edgar and Shaw 1995b). Making predictions is often necessary because dietary data do not typically exist, or at least not for the majority of species present in a fish community. Dietary predictions for individual animals will allow more accurate models of community consumption and food webs with which to answer key questions about the ecology of marine systems in relation to human impacts and management interventions, in the context of the effects of environmental variables. Nevertheless, the challenge with testing ecological predictions at large spatial scales is considerable due to logistical constraints and the great spatial variability in the natural environment (Edgar and Shaw 1995b; Peters 1991). Community-level calculations are useful, such as when assessing impacts of fishing and other threats on food web processes. At present, it is near-impossible to identify community-level patterns of consumption from integration of limited dietary data available for particular fish species, given the diversity of fishes of different body sizes at any single location, let alone across multiple locations.

A predictive diet model, *Consume*, was developed using dietary data from shallow water marine fishes sampled in Western Port, Victoria (Soler et al. 2016). Some of the findings of the study in which this model is described included: (a) prey type and mean prey size were accurately predicted for consumer fish of known species and size (77% accurate for prey type and 93% for mean prey size; prey type accuracy was calculated as the mean percentage overlap between predicted and actual prey types, and mean prey size accuracy as the correlation (r^2) between predicted and observed mean prey size); (b) when only

* Corresponding author.
E-mail addresses: german.soler@utas.edu.au (G.A. Soler), G.Edgar@utas.edu.au (G.J. Edgar), Rick.StuartSmith@utas.edu.au (R.D. Stuart-Smith), Tony.D.Smith@csiro.au (A.D.M. Smith), Russell.Thomson@utas.edu.au (R.J. Thomson).

http://dx.doi.org/10.1016/j.jembe.2016.03.017
0022-0981/

the family of the consumer, rather than the species identity, was included, the loss in accuracy of the models was small (~1%); (c) the most important predictor for prey type was the size of the consumer; and (d) the most important predictor for mean prey size was the taxonomic identity of the consumer.

In this study, the accuracy of the *Consume* model was tested across a larger spatial and taxonomic domain in order to determine the generality and broader utility of this novel means to quantify food web links in the absence of detailed dietary information. The fish stomach content dataset described by Edgar and Shaw (1995a, 1995b) was used for model training and field-testing, and encompassed 14 locations in southern Australia, plus five locations in Western Port, Victoria (Fig. 1). These locations extend over 3000 km of coast and six marine ecoregions of the world and two provinces (Spalding et al. 2007).

The most important predictors of fish diet were tested to determine if they could be generalized over large spatial scales, or whether important location-specific factors and local community composition result in idiosyncratic patterns of prey consumption which may prohibit accurate larger-scale size-based food web modelling. To examine this, the loss in accuracy was tested when predicting fish diet for the 14 locations in southern Australia using models trained only on the Western Port dataset, and therefore naïve to locally-collected dietary data, in comparison to model accuracy when trained with local data from a subset of sampled fishes. Finally, the loss in accuracy was estimated for prey type and mean prey size predictions when information on taxonomy, consumer fish size or locality (habitat, location, province and/or ecoregion) is lacking.

2. Methodology

A total of 4336 fish were sampled using seine and gill nets from 19 shallow marine locations in southern Australia, extending from Rottnest Island in Western Australia to Jervis Bay in New South Wales, including five in Western Port, Victoria (Fig. 1, Table 1S; Edgar and Shaw, 1995b). For each location, fish were caught, measured, weighed, and the stomach contents studied using consistent methods (see Edgar and Shaw (1995b)). Stomach contents were identified to the lowest taxonomic level possible, measured using a microscope graticule or Vernier callipers, and binned into 19 log-scale size-classes ranging from 0.125 mm to 64 mm (Edgar and Shaw 1995a).

For dietary predictions, the *Consume* model developed in R–Studio (R-Core-Team, 2014) in a previous study (Soler et al. 2016) was used. *Consume* has two steps, the first step involved predicting the percentage of different prey types for each individual fish. For the second step, the mean prey size was predicted for a given prey type of an individual fish. The prediction of prey type percentage for a given fish was made using a set of fish with known diets, that are most similar to the given fish. This method is akin to a *k*-nearest neighbour procedure (Barber 2011; Conway and White 2012). The prediction of mean prey size (step 2) was made using linear regression models, where the assumptions of normality were met by applying a log transformation to the mean prey size. The importance of predictors was evaluated via cross validation for both steps of the *Consume* model.

In order to estimate the effect of location in the diet predictions, province and ecoregion (Spalding et al. 2007) were included in the *Consume* model. The Western Port and southern Australia dataset was

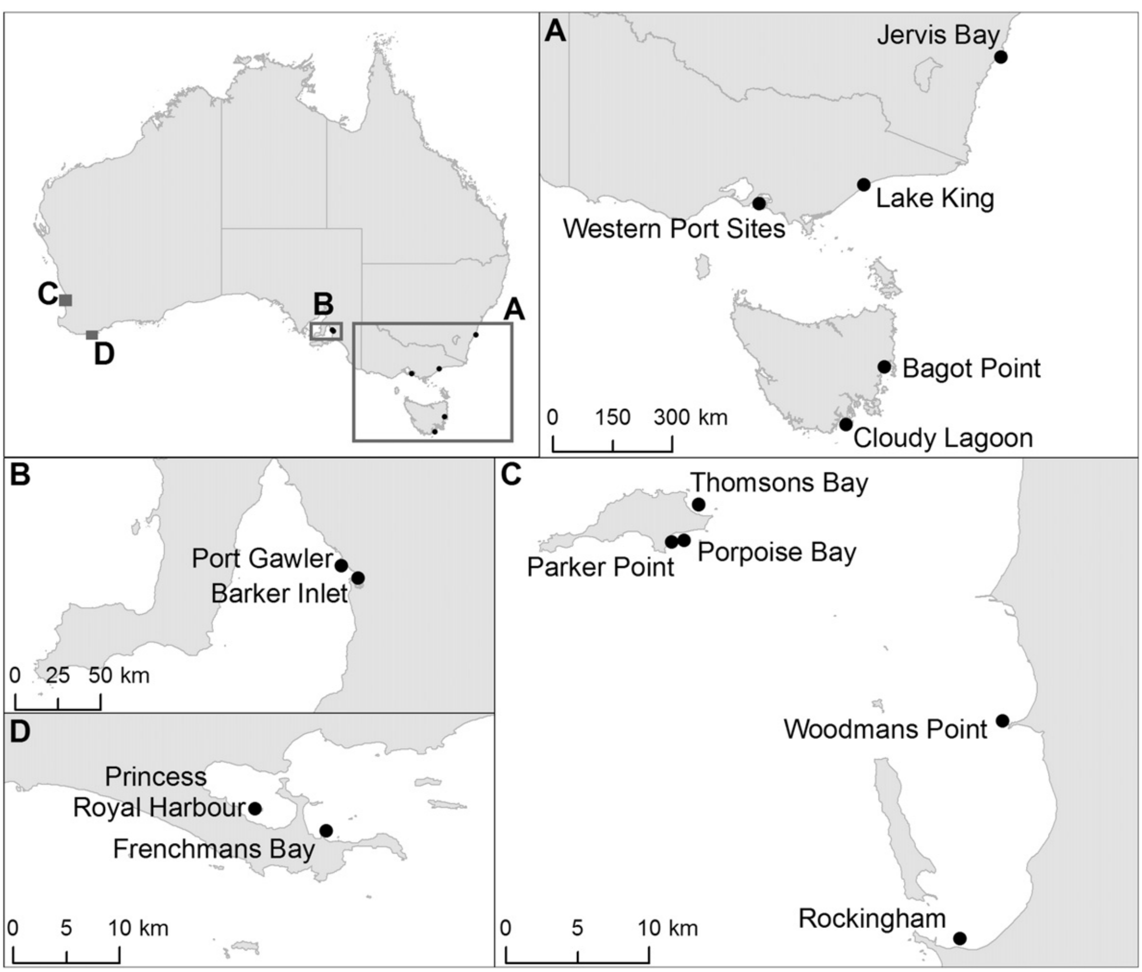

Fig. 1. Locations sampled in southern Australia.