Exploring the role of environmental variables in shaping patterns of seabed biodiversity composition

C. Roland Pitcher[1,7], Peter Lawton[2,6], Nick Ellis[1,7], Stephen J. Smith[2], Lewis S. Incze[3], Chih-Lin Wei[4], Michelle E. Greenlaw[2], Nicholas H. Wolff[3], Jessica A. Sameoto[2], Paul V.R. Snelgrove[5]

Environmental variables are increasingly used as indirect surrogates for mapping patterns of biodiversity because species survey data are scant, especially in the marine realm. This cross-program synthesis quantified the shapes and magnitude of multiple species responses to >30 environmental gradients, and the extent to which these variables were useful as predictors of biological patterns.

Methods: We collated broad scale biological survey datasets from trawl, benthic sled and grab sampling, and environmental data layers, for three large marine regions: the Great Barrier Reef, Gulf of Mexico, and Gulf of Maine. We developed and applied a novel statistical approach, "*Gradient Forest*", a modification of the tree method "Random Forest", to explore the pattern and extent of compositional changes along each environmental gradient.

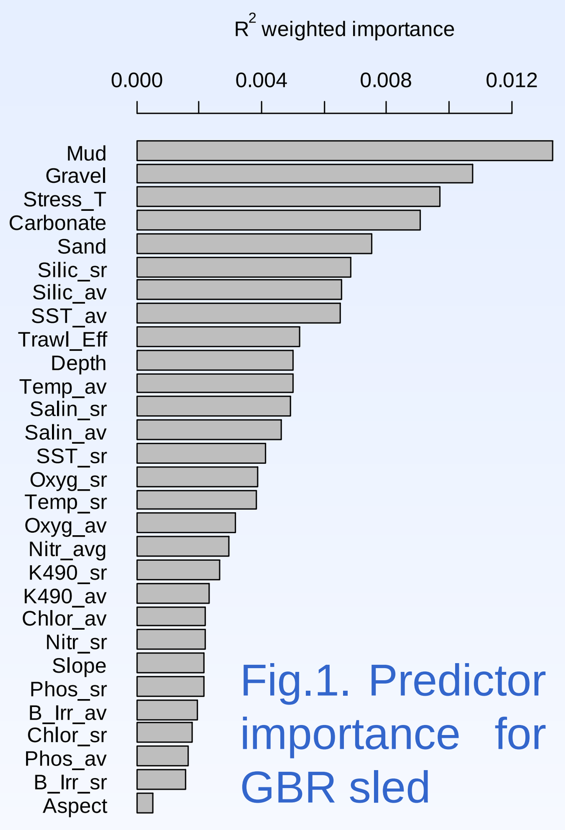

Fig.1. Predictor importance for GBR sled

Analysis: A forest of 500 trees was run for each of 100s of species in each dataset, from which predictor importance was assessed (e.g. Fig.1), and each split value and deviance improvement was extracted (Fig. 2a). The split-improvements were aggregated and standardised by data density to show where important changes occurred along the gradients (e.g. Fig.2b). Cumulative distributions of the splits show overall changes for individual species (e.g. Fig. 2c) and for the whole community (e.g. Fig.2d) in units of R^2.

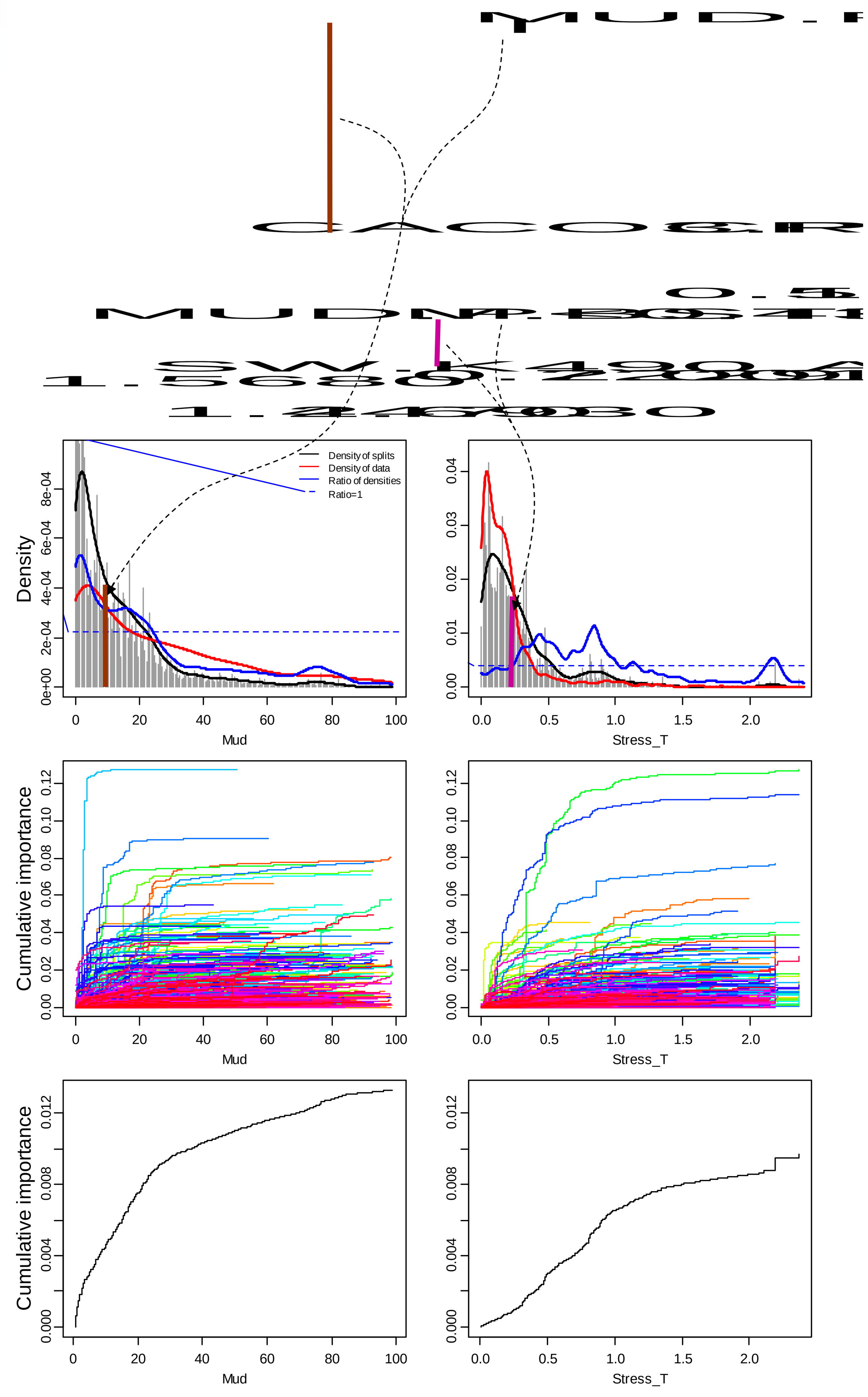

Fig.2. (a) a single tree from a forest, (b) splits density distributions, (c) species cumulative R^2 importance distributions, (d) community cumulative importance

Results: Over all three regions, the environmental variables together typically predicted ⅛ to ⅓ of the distributional variation in species abundance patterns (average R^2=0.13–0.35, max. 0.5–0.8). Important predictors differed among regions and biota, and included depth, salinity, temperature, sediment composition, and current stress (Fig.3). The shapes of responses also differed and were non-linear, often with thresholds indicative of step-changes in composition (Fig.3). The largest compositional changes were across provincial transitions and coincided with sharp thresholds on several gradients.

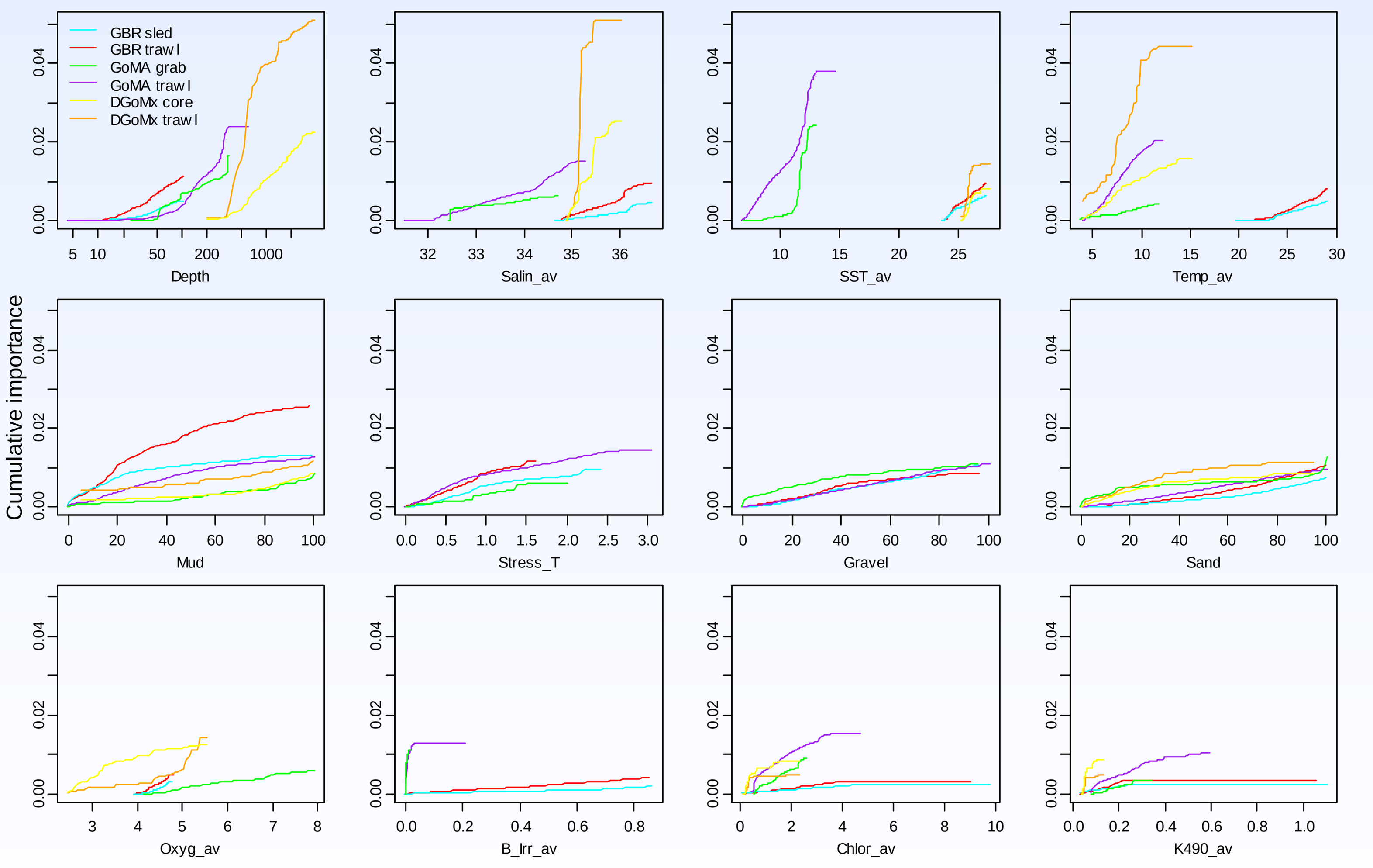

Fig. 3. Cumulative importance (R^2) distributions of selected influential predictors available for two or more regions of three regions, in order of overall importance.

Conclusions & Applications: The differing regional responses in analyses to date may indicate limited scope for extrapolating bio-physical relationships beyond source regions. Within regions, Gradient Forest provided insightful exploration of species-environment-relationships and, importantly, permits combining information from disparate datasets. The main outputs, cumulative importance distributions (e.g. Fig.4a), can be used as biologically-informed transformations of environmental layers to provide surrogate maps of expected patterns of biodiversity composition (e.g. Fig.4b).

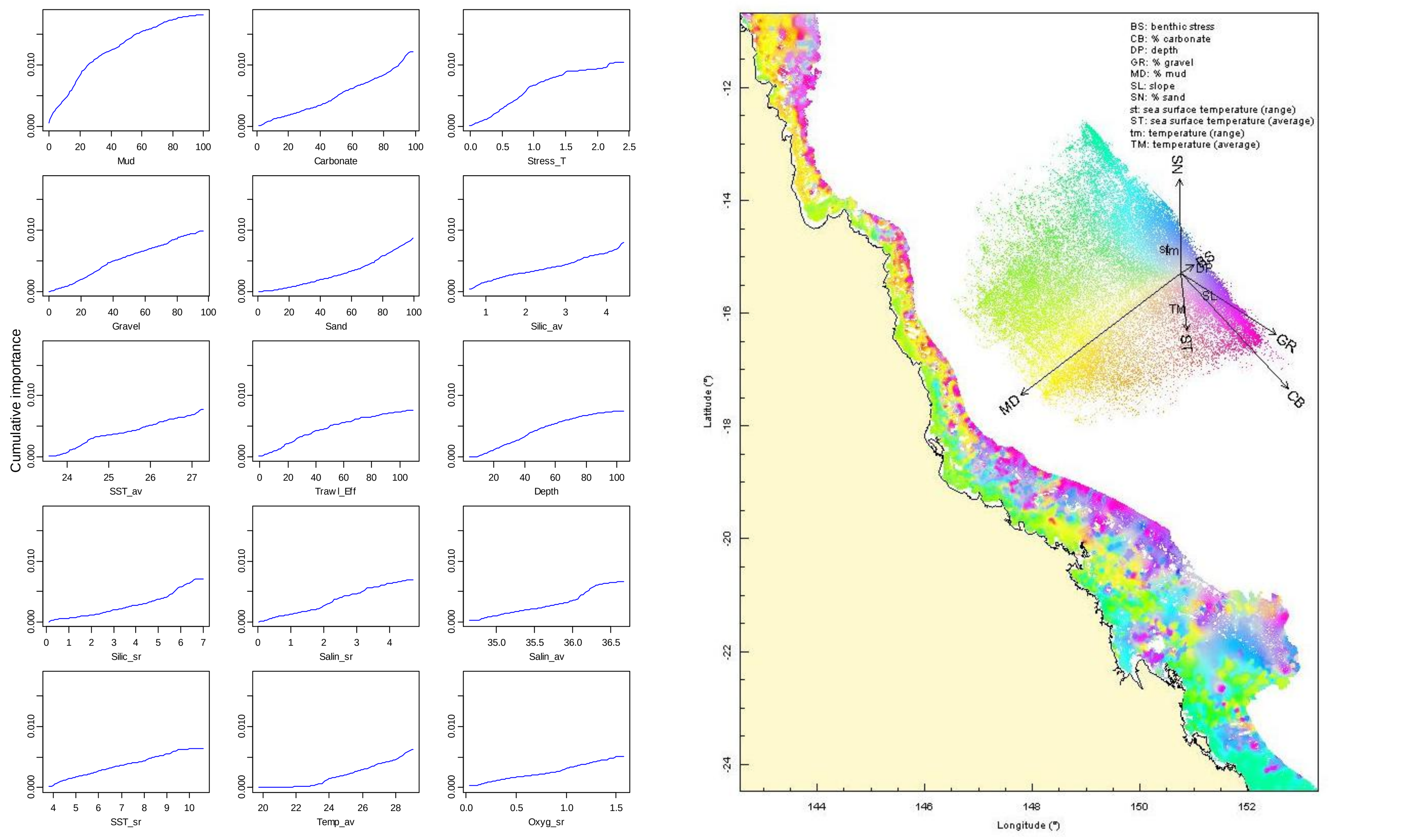

Fig.4. (a) predictor cumulative importance distributions for 15 of 29 predictors used to transform and (b) map the GBR environment space, representing the first two dimensions of expected continuous patterns of biodiversity composition.

Acknowledgements: We thank the agencies that provided data for the analyses. This work was supported by the authors' agencies and by CoML.

A DECADE OF DISCOVERY CENSUS OF MARINE LIFE 2010

1 CSIRO
2 Fisheries and Oceans Canada
3 University of Southern Maine usm
4 Texas A&M University at Galveston 1876
5 Memorial University
6 Centre for Marine Biodiversity
7 Marine Conservation Research — Prediction and Management of Australia's Marine Biodiversity

Contact: Roland.Pitcher@csiro.au