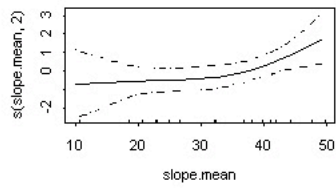**Figure 6**. Multi-regression model results for the proportion of size classes >125 mm for *Evechinus chloroticus*. Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.

Community and Functional Group-Level Indicators. The variation of species density (i.e. standard deviation) showed a curvilinear relationship with mean slope within the coastal point buffers (Figure 7). This showed a positive increase in species density variation with increasing slope, which became more steep with slopes greater than 40 degrees. Validation and cross-validation tests showed a relatively lower correspondence between tests, with an r-value of 0.6 for the validation test and an r-value of –0.011 for the cross-validation test (Figure 7). Residuals for this model showed some degree of departure from linearity (Figure 7).

Observed species richness of the epifaunal assemblages showed relationships with mean depth and mean slope, where both were negative trend lines (Figure 8). These trends were very close to linear, with some deviation from linearity for high values of slope (i.e. > 40 degrees). Validation and cross-validation tests showed r-values of 0.80 and 0.61, respectively. Residuals for this model showed some degree of departure from linearity (Figure 8).

The proportion of the number of ascidian taxa showed a relationship to the standard deviation of mean annual solar radiation, indicating a relatively lower proportion of ascidians with lower variation until a standard deviation of approximately 3.0, above which there was a near-linear increase (Figure 9). This pattern was driven by a relatively fewer number of points in areas of high variation, which accounts for the increased error in the model along this portion of the axis. Validation and cross-validation tests indicated a r value of 0.63 and –0.68, respectively (Figure 9). There was a slight departure from linearity in the quantile plot of residuals.

GRASP: StdevD

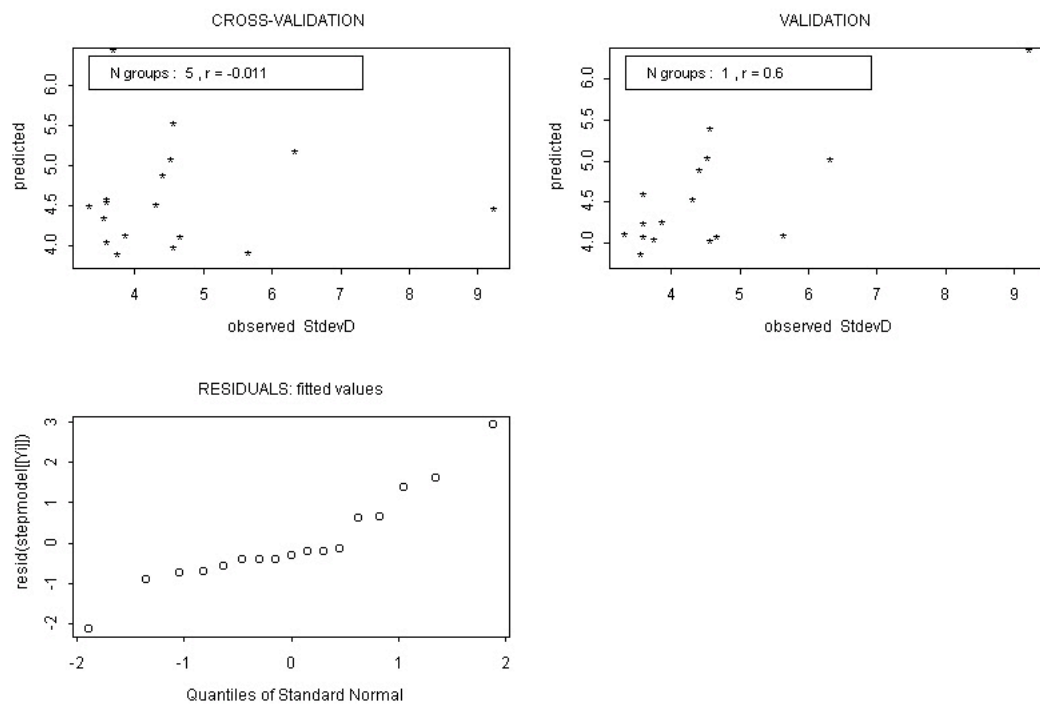

formula = YYY$StdevD ~ s(slope.mean.2), family = gaussian

CROSS-VALIDATION

VALIDATION

N groups : 5 , r = -0.011

N groups : 1 , r = 0.6

RESIDUALS: fitted values

Figure 7. Multi-regression model results for the average variation (i.e. standard deviation) of species density per 0.25 m^2 quadrat. Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.
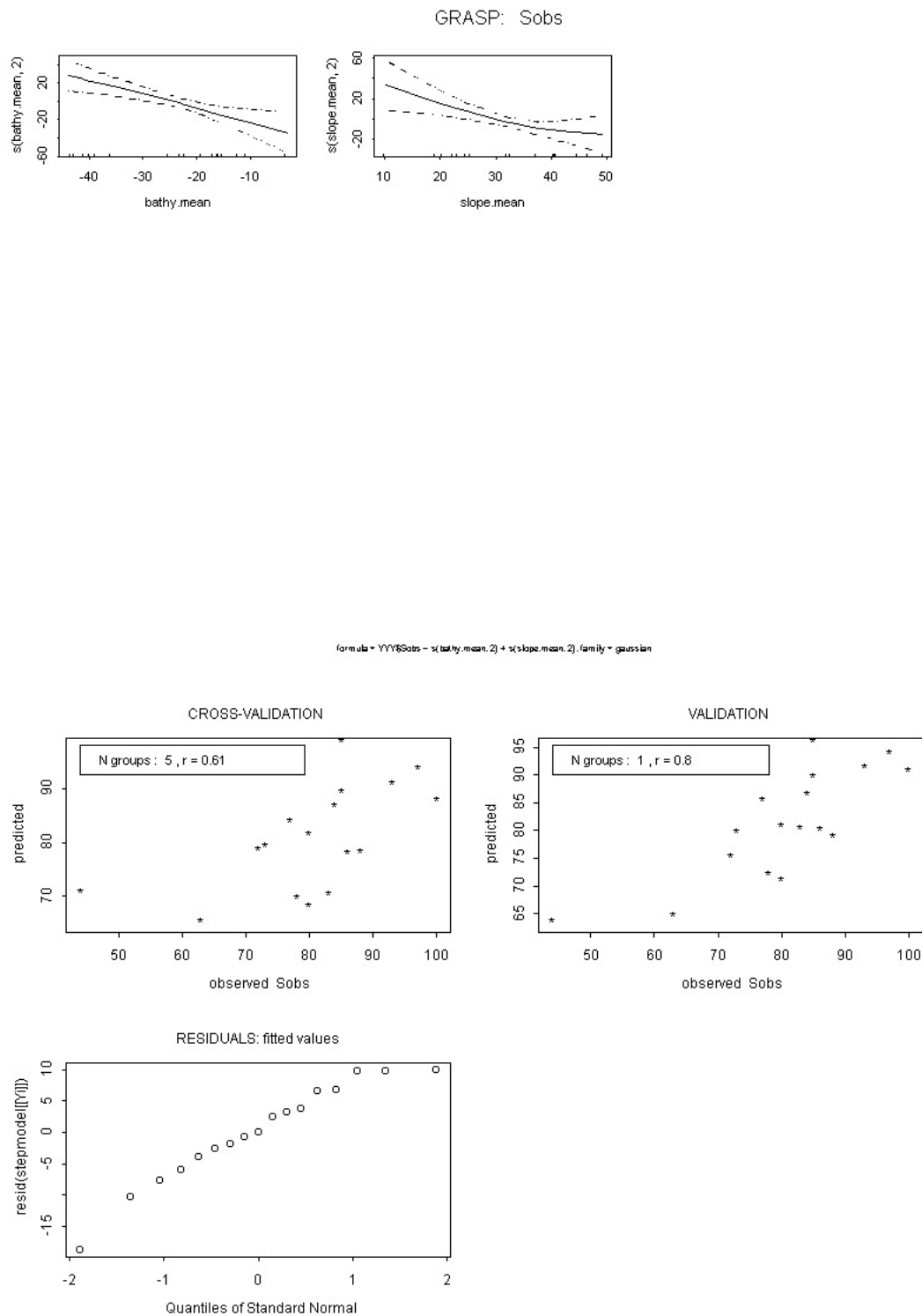
Figure 8. Multi-regression model results for the average variation (i.e. standard deviation) of species density per 0.25 m² quadrat. Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.
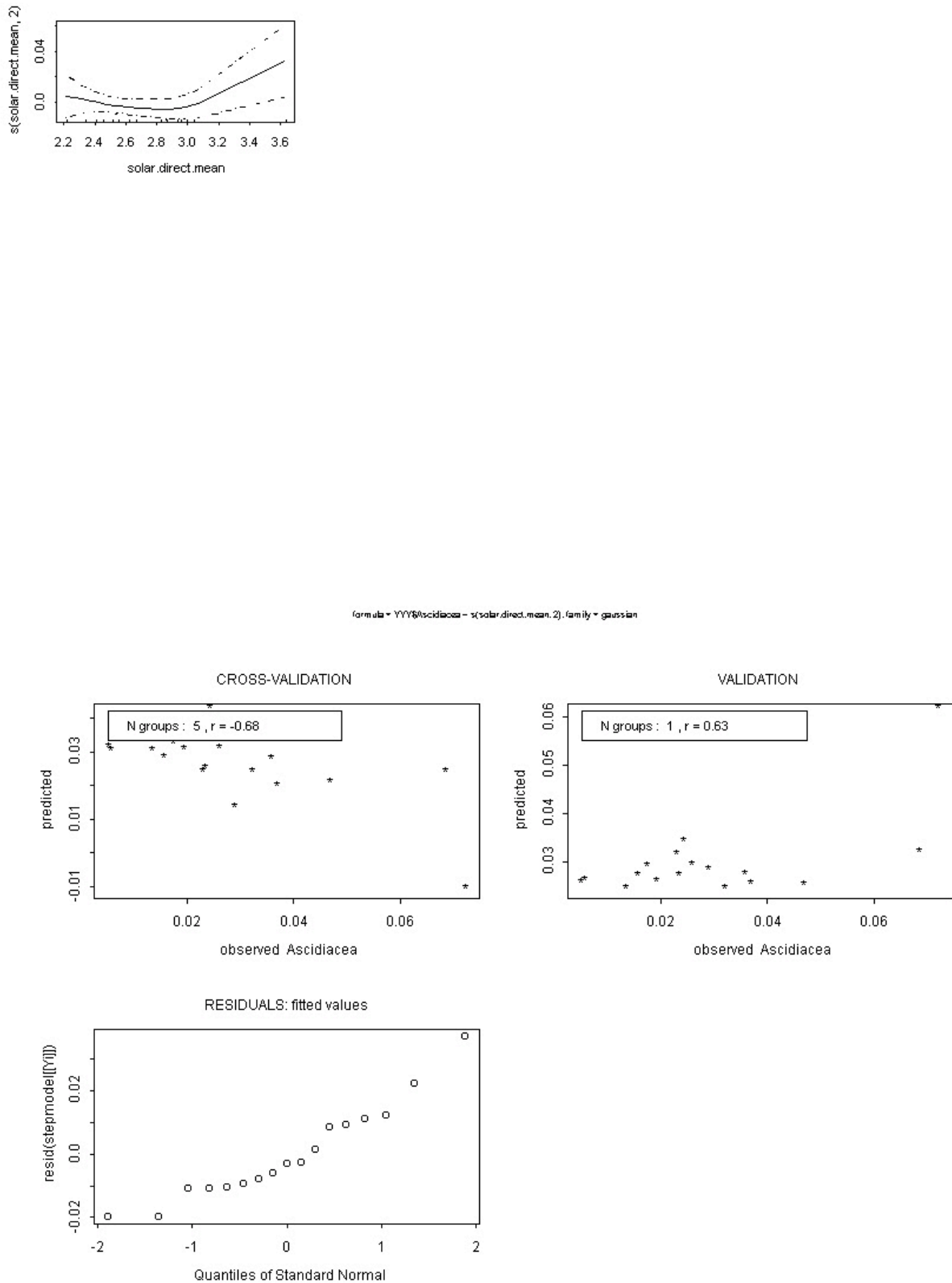
GRASP: Ascidiacea

Figure 9. Multi-regression model results for the relative number of ascidian species as a proportion of the total number of species. Lines refer to the predictions for environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.

Species-Level Indicators.   The relative abundance of butterfly perch (*Caesioperca lepidoptera*) showed a curvilinear relationship with global solar radiation, where sites of lower solar radiation (i.e. < 1700 Wh·m^{-2}·day^{-1}) tended to have higher relative abundance (Figure 10).   This trend line levelled off at about 1800 Wh·m^{-2}·day^{-1}, where there were comparatively lower numbers of butterfly perch. Validation and cross-validation tests for this model showed r-values of 0.63 and 0.36, respectively, and residuals showed some departure from linearity (Figure 10).

Relative abundance of scarlet wrasse (*Pseudolabrus miles*) showed a relationship with mean significant wave height (H$_{sig}$) and water column mixing (Figure 11). Both trend lines were curvilinear, with the relationship with H$_{sig}$ suggesting a unimodal function with a maximum abundance at wave heights of 0.9 m. The trendline with respect to water column mixing showed a linear increasing trend to about 0.7, where there was a sharp increase in the relative abundance of scarlet wrasse to mixing values of about 0.8. The validation and cross-validation tests for this model resulted in r-values of 0.79 and 0.58, respectively (Figure 11). Fitted residuals for this model showed a near linear trend, with a slight departure at standard normal quantiles above 1.5 (Figure 11).

One physical environmental variable, surface salinity, was related to the relative abundance of blue cod (*Parapercis colias*). This relationship was curvilinear with a local maximum at surface salinities of approximately 27 psu, where the relative abundance of blue cod decreased sharply (Figure 12). There was a relatively larger degree of error associated with this trend line at surface salinities < 20 psu, with few observation points in areas of lower salinity. The validation and cross-validation tests for this model had r-values of 0.65 and 0.46, respectively, and plotted residuals showed some departure from linearity (Figure 12).

GRASP: SCL
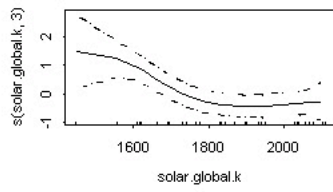
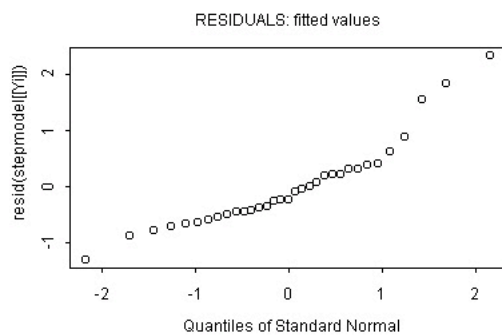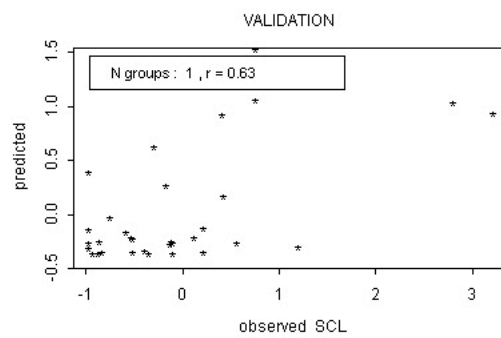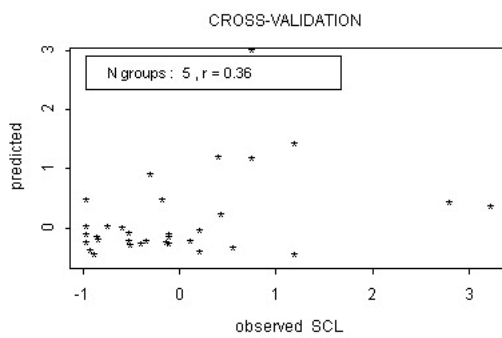formula = YYY$SCL ~ s(solar.global.k, 3), family = gaussian

CROSS-VALIDATION

VALIDATION

N groups : 5 , r = 0.36

N groups : 1 , r = 0.63

RESIDUALS: fitted values

Figure 10. Multi-regression model results for the relative abundance of butterfly perch (*Caesioperca lepidoptera*). Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.
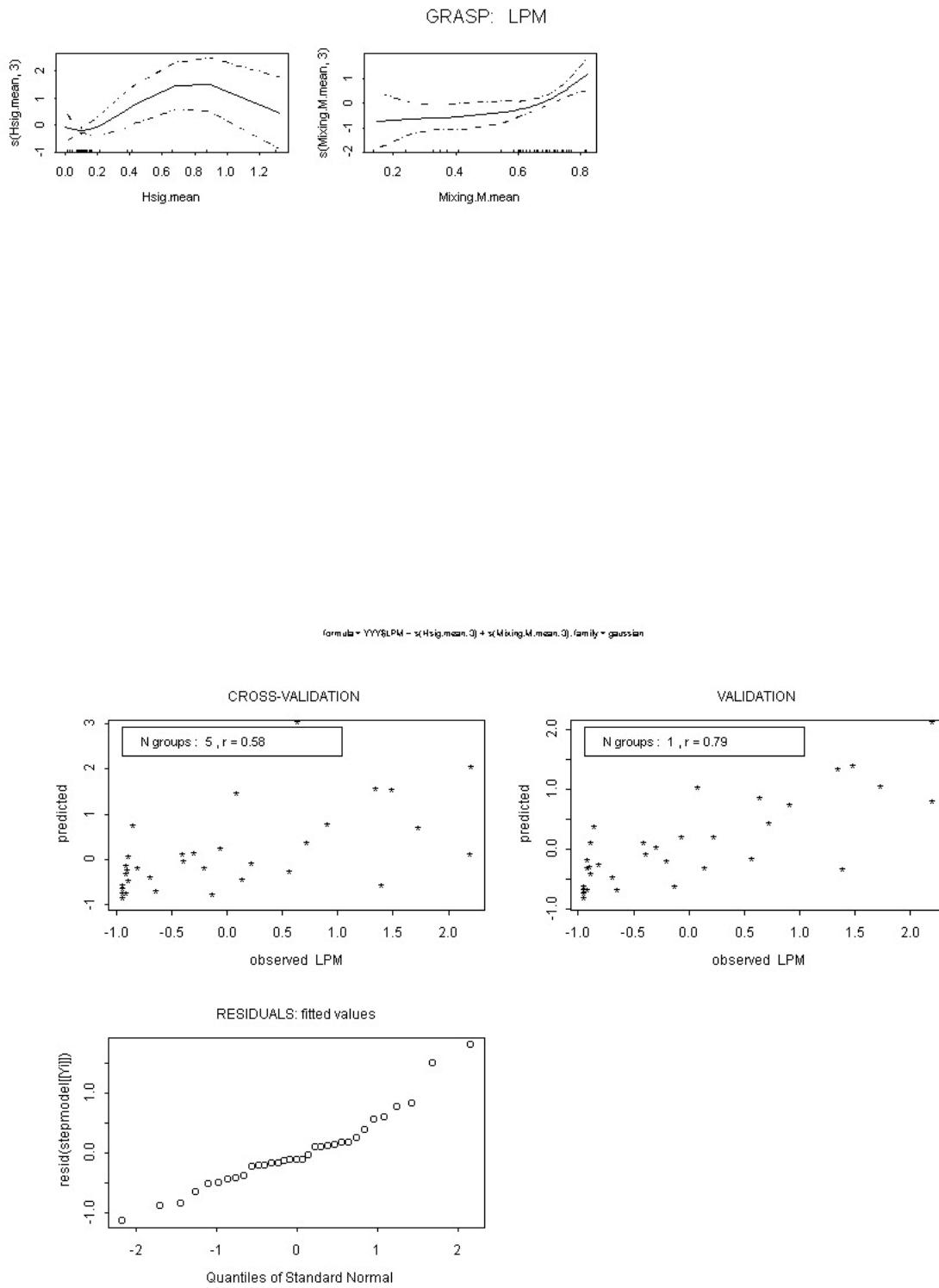
Figure 11. Multi-regression model results for the relative abundance of scarlet wrasse (*Pseudolabrus miles*). Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.

GRASP: PPC

Figure 12. Multi-regression model results for the relative abundance of blue cod (*Parapercis colias*). Lines refer to the predictions for individual physical environmental variables and the dotted lines refer to the confidence intervals. Whisker plots at the bottom of each graph refer to the location of actual data points. The lower portion of the figure shows the results of the cross-validation, validation and residuals for the model.
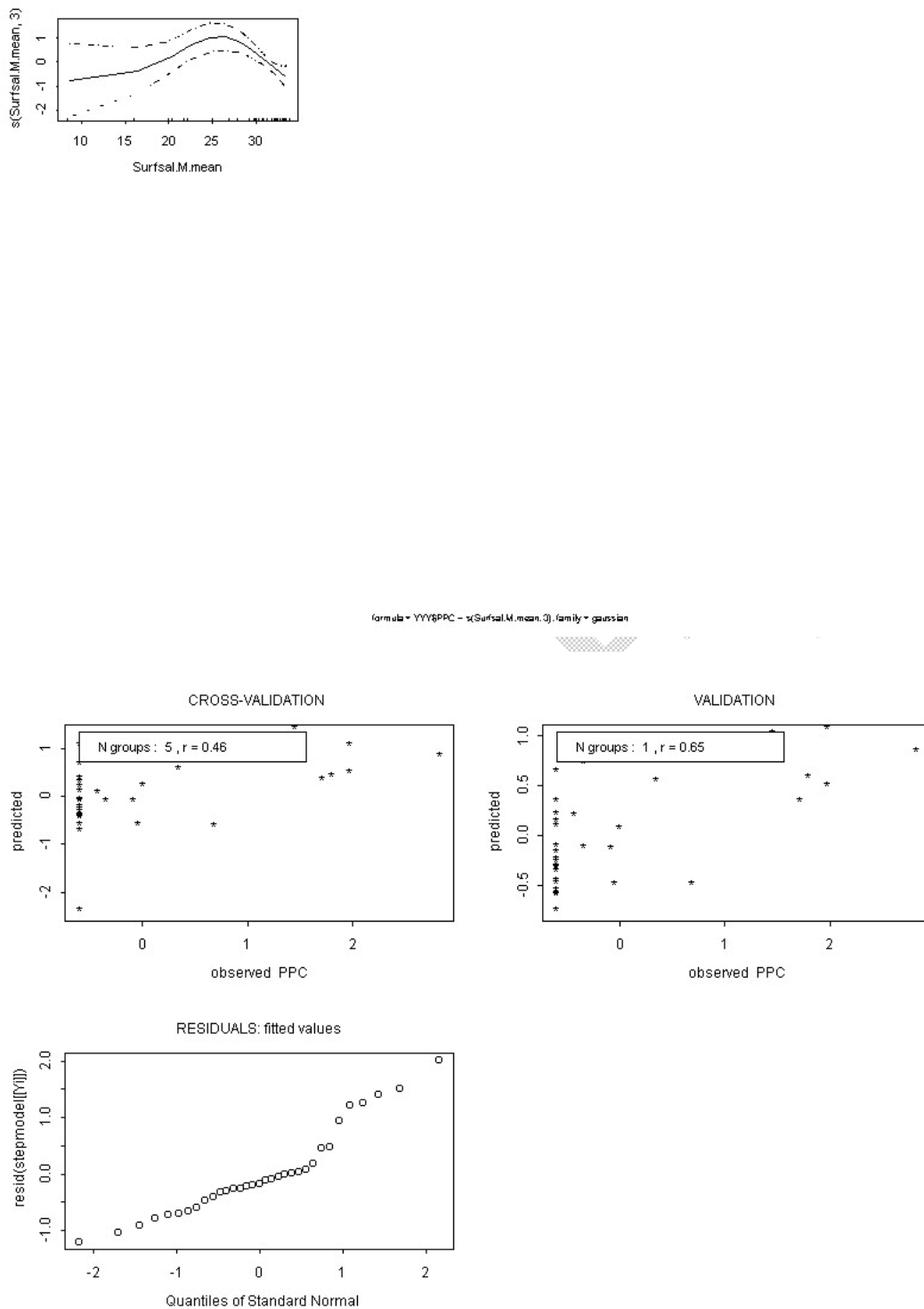
Ground-truthing Results

Population-level Indicators. Predictions of aspects of the size structure of sea urchin population and observed values from the ground-truthing showed a correlation of 0.214 with orthogonal regression (Figure 13A). This regression had a slope of approximately 9.72, which was considerably steeper than expected relationship (i.e. slope = 1). The model-driven regression showed an overlap of the 1:1 correlation line, indicating that data generally conformed to the model (t-ratio = -0.02, P-value = 0.988), with outliers for sites with larger sea urchins (Figure 13A).

The orthogonal fit for data of sea urchins >125 mm showed a correlation of 0.083 with a slope of over 21 (Figure 13B). This was largely due to two points with considerable leverage, one where there was a higher proportion than predicted (i.e. about 5-fold higher, site #5) and one site where predicted proportion of larger-sized urchins was not observed in the ground-truthing data (i.e. site #41). Linear model fits of constant slope of 1 showed that the model generally predicted a higher proportion of larger sea urchins (i.e. > 125) than observed (Figure 13B). The 1:1 correlation line was well within the confidence limits for this regression, although about half of the sites did not conform to the model (t-ratio = 0.85, P-value 0.415).

For the proportion of sea urchins < 70 mm in test diameter, the orthogonal regression showed a correlation of –0.316, where the negative slope indicates that there was generally a higher proportion of smaller urchins present than predicted from the multi-regression model (Figure 13C). For this test, one site had considerable leverage, which predicted a comparatively higher proportion of smaller sea urchins than were observed in the ground-truthing data set (i.e. site #37, Figure 13C). The model-driven regression showed that the 1:1 regression line was located within the confidence levels of the line fitted through the data with a slope of 1 (t-ratio = 0.93, P-value = 0.373). A number of sites did not conform to this model, particularly those where there were a smaller proportion of urchins < 70 mm predicted (e.g. sites #28, 30 and 41)(Figure 13C).
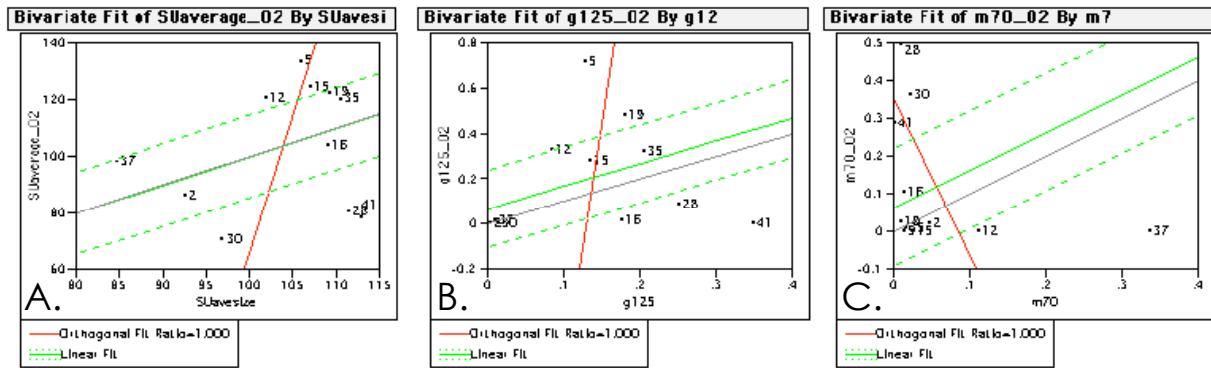
Figure 13. Relationships between predicted (A) sea urchin average size, (B) proportion of large (i.e. > 125 mm), and (C) small (i.e. < 70 mm) sea urchins. Red lines indicate an orthogonal fit of the data points with equal variation. Solid green lines indicate a regression with a slope of 1 and the dotted green lines are the standard errors of the regression. Grey lines indicate a regression line with a slope of 1 that is directly proportional to both axes. See Methodology and Results for explanation of statistical techniques and results for individual tests.

Community and Functional Group-Level Indicators. Orthogonal regression of predictions of variation of species density of the sessile suspension-feeding invertebrates showed a correlation of 0.489 with a slope of 20.87 (Figure 14A). This trend was largely driven by one site where there was a 2-fold higher variation of species density observed than predicted (i.e. site #24). Model-driven regressions of a slope of 1 showed that the majority of sites were outside the confidence levels of the regression. The t-test for differences in the intercepts of these two regressions had a t-ratio of 2.15 with a P-value of 0.063, indicated by the the 1:1 line lying on the inside edge of the confidence levels (Figure 14A).

Predictions of species richness showed a correlation of –0.454 with the observed species richness from the ground-truthing data set, under-predicting the number of species (Figure 14B). Several sites were important in driving this relationship, in particular site #9 and 24 (under-predicted) and #4 and 20 (over-predicted). Comparisons between the models with a slope of 1 showed significant differences, where the observed species richness tended to be under predicted by the multi-regression models (t-ratio = 4.3, P-value = 0.0026)(Figure 14B).

Modelled predictions of the proportion of ascidians showed a correlation of 0.070 and a slope of 142.45, indicating a weak correspondence between predicted and observed levels for this taxa (Figure 14C). The model-driven regression with a slope of 1 showed highly significant differences between the 1:1 trend line and those fitted through the data (t-ratio =

–46.99, P-value = <0.0001). Five (5) of the 9 sites with predicted values were within the confidence bands of the model-driven regression (Figure 14C).
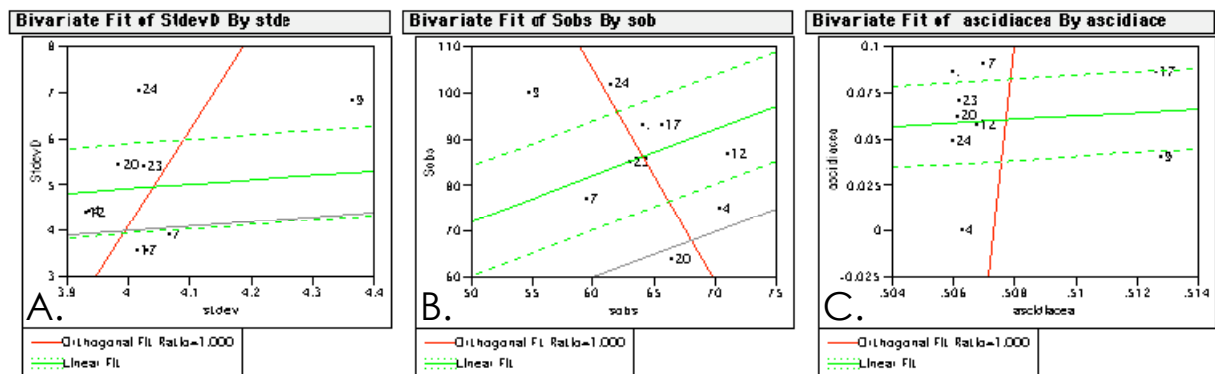

Figure 14. Relationships between (A) the standard deviation of species density of epifaunal suspension-feeders; (B) observed species diversity (i.e. S_{obs}); and (C) proportion of ascidian species of the total number of species. Red lines indicate an orthogonal fit of the data points with equal variation. Solid green lines indicate a regression with a slope of 1 and the dotted green lines are the standard errors of the regression. Grey lines indicate a regression with a slope of 1 directly proportional to both axes. See Methodology and Results for explanation of statistical techniques and results for individual tests.

Species-Level Indicators. The predicted relative abundance of butterfly perch showed a correlation of –0.051 with the observed relative abundance in the ground-truthing data set (Figure 15A). This trend was largely due to about 1/3 of the total number of sites having larger numbers of butterfly perch than predicted by the multi-regression model and a few sites over-predicting the relative abundance (Figure 15A). The model-driven regression showed few sites within the confidence bands, although the 1:1 regression line was located just inside the confidence levels of the model fitted through the data (t-ratio = –0.31, P-value = 0.756).

Orthogonal regression between predicted and observed data of the relative abundance of scarlet wrasse showed a correlation of –0.198, having a similar trend to butterfly perch, with higher observed abundances than predicted at about 1/3 of the sites and a few sites that over-predicted the abundance levels (Figure 15B). Constant-slope models showed few sites located within the confidence levels, largely for sites with lower predicted and lower observed abundances. The 1:1 model was located just inside the confidence levels of the model fitted to the data (t-ratio = 1.38, P-value = 0.176).

The orthogonal regression for predicted and observed relative abundance for blue cod was largely driven by a single point (site #51), with an overall correlation of 0.361 and a slope of 5.707 (Figure 15C). Model-driven regressions showed similarities between the directly

proportional model (i.e. 1:1) and the regression through the data, with a t-ratio of-0.31 and a p-value of 0.75 (Figure 15C), indicating that there is no statistical difference between the intercepts of these regressions. A number of sites predicted to have larger relative number of blue cod than observed (e.g. sites #15, 26, 38 and 50), indicating that the predictive model tended to over-predict for locations with higher relative abundance (with the exception of site #51).
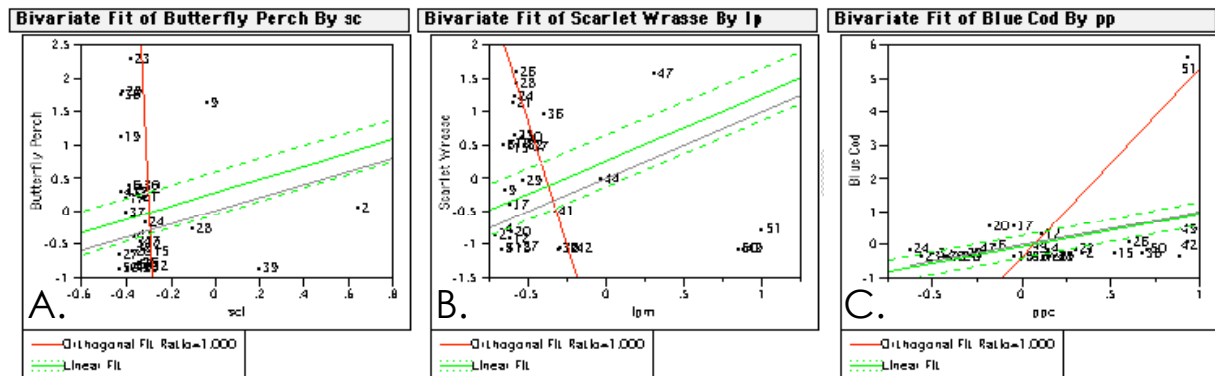

Figure 15. Relationships between predicted abundance of (A) butterfly perch, (B) scarlet wrasse, and (C) blue cod. Red lines indicate an orthogonal fit of the data points with equal variation. Solid green lines indicate a regression with a slope of 1 and the dotted green lines are the standard errors of the regression. Grey lines indicate a regression with a slope of 1 directly proportional to both axes. See Methodology and Results for explanation of statistical techniques and results for individual tests.

Discussion

Multi-regression Modelling of Biological Parameters. Results of the multi-regression modelling indicate that several of the physical environmental layers were consistently selected as correlates for different biological response variables. For example, mean annual solar radiation and significant wave height (H_{sig}) were selected for all 3 sea urchin models screened (Figures 4-6). Slope was selected as an important variable for 2 of the 3 epifaunal invertebrate models selected to model (i.e. Figures 7-10). Models of fish abundance produced a relatively broader selection of variables, including solar radiation, significant wave height, mixing, and surface salinity (Figures 10 -12). This suggests that there is more consistency in the relevance of certain physical variables within a level of biological organisation (i.e. population- or community-levels) than across different levels. The larger number of physical variables selected for reef fish abundance seems to suggest that the species-level of biological organisation has less consistency when comparing different species. Although it is difficult to generalise the importance of individual physical environmental parameters determining

species level of biological organisation, information on physical parameters which are correlated with the abundance of individual species can still be informative for understanding how species assemblages may be constrained. A more detailed analysis of different species and their physical correlates would be required to better understand potential habitat requirements for an entire species assembly.

Each suite of models differed with respect to the cross-validation results, where r-values for sea urchin models ranged between 0.17 and 0.49; epifaunal invertebrate models between 0.01 and 0.61; and fish models ranging between 0.36 and 0.58. These results suggest that there is no general tendency for models to be more accurate for different levels of biological organisation, based on a similar suite of predictor variables for the study region. Differences in the number of sites with biological information and the overlap with regions of continuous physical environmental data available may account for some of the observed variation between the different data sets. For example, data to model parameters related to fish abundance was based on information from 61 sites where there was corresponding physical data, 18 sites for sea urchins and 17 sites for epifaunal invertebrates. As fish models had more information available to model with potential physical environmental correlates, the relatively higher r-values from the cross-validation tests may be partially attributable to this factor.

Ground-Truthing of Predicted Results. The relationship between biological parameters predicted through multi-regression statistics and values observed during the ground-truthing sampling showed a large degree of variation depending on the response variable. Some models, such as average sea urchin size, proportion of small sea urchins (i.e. < 70 mm) and blue cod abundance performed relatively well according to the model-driven test for predicted versus observed values (Figures 13A,B and 15C). Other models did not have a strong relationship with the 1:1 predictions, such as the standard deviation of species density, scarlet wrasse and butterfly perch (Figures 14A and 15A,B), while the remaining models were significantly different from an expected 1:1 correlation.

Orthogonal regressions provided information on how individual data from models corresponded to those observed in the ground-truthing survey. These correlations were typically low (i.e. r-values <0.2), with the exception of the variation of species density (Figure 14A) and the abundance of blue cod (Figure 15C). The slopes of these correlations were often much greater than the expected 1, and in some instances, slopes were negative. This indicates that models commonly over-predicted or under-predicted for individual sites and it

was common that some sites had considerable leverage in the direction and inclination of the correlations (e.g. site #5 and #41, Figure 13B). For dynamic elements of ecosystems, the ability to make direct correlations for data-driven models may prove difficult. Results of the orthogonal regression also suggest that the utility of predictive modelling to make reliable predictions of specific parameters does not depend on the level of biological organisation (i.e. between population, species-, and community-levels).

An overall test between the accuracy of the original multi-regression model used to make the predictions (measured by the validation and cross-validation tests) and the correlation between the predicted and observed indicates that a large amount of unexplained variation exists within the performance of individual multi-regression models and their accuracy in ground-truthing in the field (Figure 16). This linear regression was not significant (r^2 =0.25, P-value = 0.205) and an orthogonal regression correlation had an r-value of –0.502. This tendency indicates that models that were less accurate had a generally lower correlation between predicted and observed in the ground-truthing. Although the regression was non-significant, this approach may prove useful in determining the reliability of a suite of models of different biological attributes in the development of a classification. That is, descriptions of biological variation according to physical environmental parameters that have less explanatory power may be used more cautiously in decision-making and/or used to focus research questions or sampling effort to be able to improve the overall performance of a classification for different levels of biological organisation.

Figure 16. Correlation between the difference between validation and cross-validation tests of the Generalised Additive Model and the absolute value of the correlation coefficient of the orthogonal regression between predicted and observed values from the ground-truthing tests. Red line represents a linear fit through the data (r^2 = 0.25, F-ratio=2.017, P-value 0.205) and the green line represents an orthogonal regression (correlation coefficient = -0.502). Numbers refer to the individual tests (1 = sea urchin mean size; 2 = sea urchins < 70 mm; 3 = sea urchins > 125 mm; 4 = variation of epifaunal species density; 5 = observed epifauna species richness; 6 = proportion of ascidian species; 7 = butterfly perch abundance; 8 = scarlet wrasse abundance; 9 = blue cod abundance).
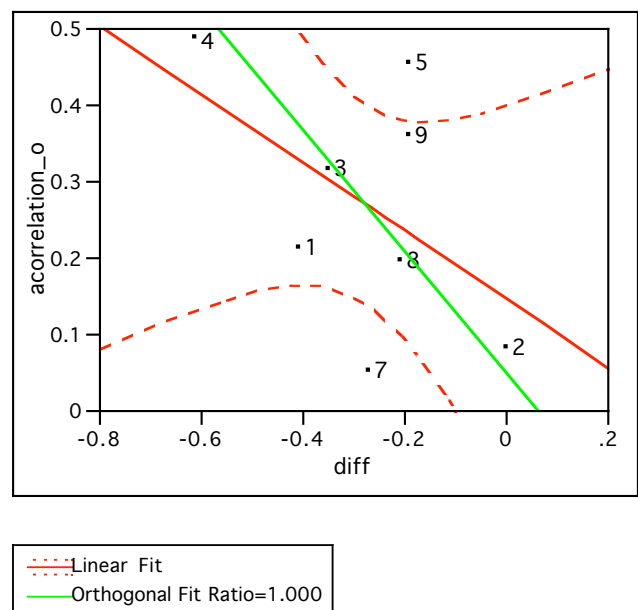
Part 2. Multivariate Classification of Coastal Environment Types and Test of
   Classification Groupings Using Biological Data

Methodology & Approach

Construction of a Physical-Based Classification of Marine Environment Types. Based on the
results of the multi-regression modelling (Part 1), physical variables available had some
correlation with biological parameters, with the exception of PEA and EFD. With these
exceptions, and additional layers of wave climate not used in the statistical modelling, were
used for the construction of the classification. Information from the physical data layers of
depth, surface salinity, temperature, wave climate, and light were extracted within 100 m
radius buffers along the coastal fringe of the Fiordland GIS (Wing et al. 2003, 2004). This
information was reduced to coastal segments of 200 m (i.e. within the buffers) where there
was overlapping information of all available data layers, including those with coverage during
CTD surveys. It should be noted that a combination of individual CTD survey coverage and
additional topographical/physical data (e.g. wave climatologies) would have provided more
extensive cover of coastal areas, however it was considered that this approach would give
unnecessary weight to a single CTD survey. This constrained the construction of the
classification to areas where there was consistent quantity and quality of physical
environmental data.

       The resultant number of cells (or coastal buffers) that filled this criteria was 3,851,
where a non-hierarchical clustering technique was then used to reduce the number of coastal
segments before using an average linkage hierarchical classification, following methodology
used in the development of the Land Environment Classification (LENZ, Leathwick, et al.
2003) and the Marine Environment Classification (Snelder, et al. 2004). For this, a non-
hierarchical clustering algorithm, the ALOB routine, was used in the computer program
PATN version 3.03 (Belbin 1995). This reduced the total number of coastal buffers to 270
groups, where the means (i.e. centroids) and standard deviations of the coastal segments were
to serve as the basis of the classification. The relationships between the original coastal
segments (i.e. buffers) with the resultant non-hierarchical clustering algorithm and the average
linkage hierarchical clustering were maintained throughout the process, to be able to map the
results of the classification in geographic space.

       The combination of means and variation from the coastal buffers from the 12 base
physical variables (including depth, slope, 2 light statistics, 4 wave statistics, and 4 attributes
of average water column conditions) were used to form a hierarchical classification using a

flexible UPGMA sorting strategy implemented in PATN. More physical variables were used than the modelling of biological parameters to maximise the ability to distinguish different physical habitat types. The Gower metric was used as a distance measure with a β-value of –0.1 was used in the clustering algorithm, which slightly dilates data space (Legendre and Legendre 1998). Similar methods were used in the construction of LENZ and the MEC (Leathwick et al. 2004, Snelder et al. 2005). Physical variables were not transformed and no weighting was given.

Once producing the hierarchical classification of the 270 groups, the topology of the dendrogram was explored in terms of the average Gower distance and the number of groups (Figure 17). This showed that there was a relatively small change in Gower distance from 270 to 150 groups, and similarly from 150 groups to 80 groups. For groupings smaller than 50, there was a marked increase in the average Gower distance between groups, following a power function ($y=0.911*x^{(-0.537)}$, $r^2=0.963$).
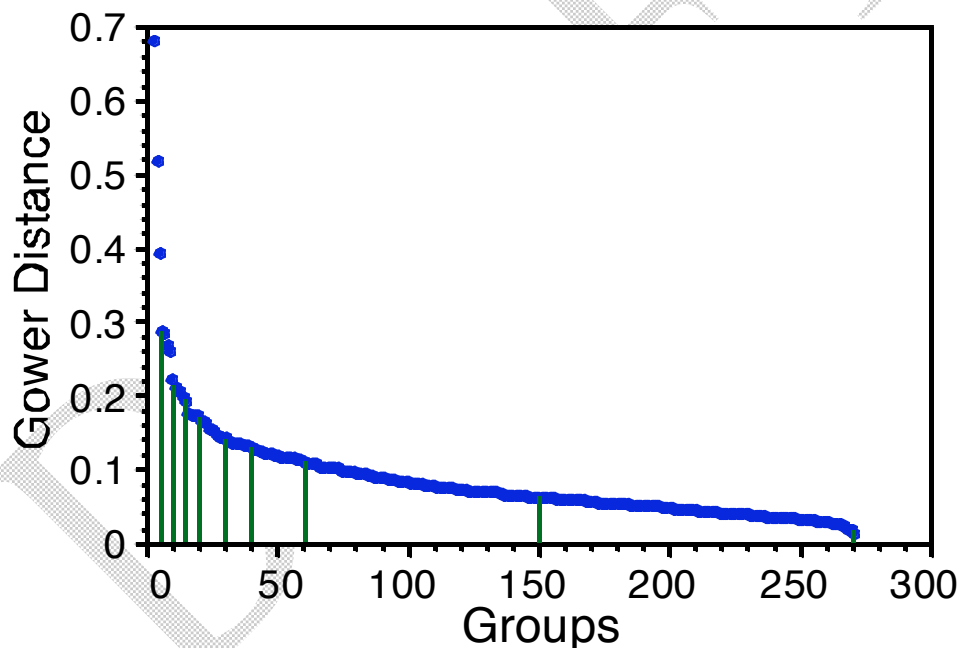

Figure 17. Relationship between environmental similarity and the number of groups in the dendrogram. Vertical lines show areas of the dendrogram that were used for comparisons of biological similarity.

Based on this pattern, relationships between group levels at 270, 150, 80, 60, 40, 30, 20, 15, 10, 7, and 5 groups were chosen to represent variation along the continuum of environmental similarity and the number of groups. As the environmental distance changes more rapidly at higher levels in the dendrogram, a disproportionate number of groupings between 40 and 5 were selected. This analysis illustrates that by using this method