Water for a Healthy Country

Ecosystem states of the Coorong: An ecosystem response model

Method development and sensitivity analyses

Rebecca E. Lester & Peter G. Fairweather

DRAFT

April 2009

Water for a Healthy Country

Ecosystem states of the Coorong: An ecosystem response model

Method development and sensitivity analyses

Rebecca E. Lester & Peter G. Fairweather

DRAFT

April 2009

The Water for a Healthy Country National Research Flagship is a research partnership between CSIRO, state and Australian governments, private and public industry and other research providers. The Flagship aims to achieve a tenfold increase in the economic, social and environmental benefits from water by 2025.

The Australian Government, through the Collaboration Fund, provides $97M over seven years to the National Research Flagships to further enhance collaboration between CSIRO, Australian universities and other publicly funded research agencies, enabling the skills of the wider research community to be applied to the major national challenges targeted by the Flagships initiative.

DISCLAIMER

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

For more information about Water for a Healthy Country Flagship or the National Research Flagship Initiative visit www.csiro.au.

Foreword

The environmental assets of the Coorong, Lower Lakes and Murray Mouth (CLLAMM) region in South Australia are currently under threat as a result of ongoing changes in the hydrological regime of the River Murray, at the end of the Murray-Darling Basin. While a number of initiatives are underway to halt or reverse this environmental decline, rehabilitation efforts are hampered by the lack of knowledge about the links between flows and ecological responses in the system.

The CLLAMM program is a collaborative research effort that aims to produce a decision-support framework for environmental flow management for the CLLAMM region. This involves research to understand the links between the key ecosystem drivers for the region (such as water level and salinity) and key ecological processes (generation of bird habitat, fish recruitment, etc). A second step involves the development of tools to predict how ecological communities will respond to manipulations of the "management levers" for environmental flows in the region. These levers include flow releases from upstream reservoirs, the Lower Lakes barrages, and the Upper South-East Drainage scheme, and dredging of the Murray Mouth. The framework aims to evaluate the environmental trade-offs for different scenarios of manipulation of management levers, as well as different future climate scenarios for the Murray-Darling Basin.

One of the most challenging tasks in the development of the framework is predicting the response of ecological communities to future changes in environmental conditions in the CLLAMM region. The CLLAMMecology Research Cluster is a partnership between CSIRO, the University of Adelaide, Flinders University and SARDI Aquatic Sciences that is supported through CSIRO's Flagship Collaboration Fund. CLLAMMecology brings together a range in skills in theoretical and applied ecology with the aim to produce a new generation of ecological response models for the CLLAMM region.

This report is part of a series summarising the output from the CLLAMMecology Research Cluster. Previous reports and additional information about the program can be found at http://www.csiro.au/partnerships/CLLAMMecologyCluster.html

Table of Contents

Acknowledgements

Executive Summary

The Coorong is an internationally-important wetland that has undergone significant degradation, particularly over the last five to ten years. The Coorong is the estuary for Australia's largest river basin, the Murray-Darling, and is of ecological significance due to the range of biota that it supports and the numbers of waterbirds that use it.

The recent degradation has led to efforts to improve the ecological understanding and management of this system. CLLAMMecology was a CSIRO Collaboration Fund Research Cluster designed to assist, particularly in the development of ecological understanding. One of the four themes within CLLAMMecology, CLLAMM Futures, was tasked with developing an ecological response model for the Coorong in order to predict the ecological response of the system to a range of potential futures, including both climate change and a variety of possible management actions. This report provides detail into the methodological development that occurred during the construction of the model, a description of the model itself and the resultant ecosystem states, and a range of model evaluation and sensitivity analyses that were conducted to understand the various limitations and sensitivities of the developed model.

The data available for the Coorong was a disparate mixture, including quantitative, semi-quantitative and qualitative data, collected using a variety of methods and at different temporal and spatial scales. The imperfect data coverage of some areas (e.g. the role groundwater and nutrients) led to a statistical approach, rather than an attempt to find deterministic relationships. A review of potentially-useful modelling techniques indicated that a state-and-transition model built using a combination of techniques to identify the ecosystem states and their transitions (including classification and regression trees (CART), multiple regression trees and structural equation modelling) was a promising option.

Two data sets were collated, a biological data set and an environmental (physico-chemical) data set. The biological data set consisted of bird, fisheries-independent fish and invertebrate abundances, catch per unit effort data from the local commercial fishery, and coverage data for the dominant macrophyte, *Ruppia tuberosa*. The environmental data set included meteorological and water quality data, along with modelled flows, water levels, depths and salinities. Variables within the environmental data set were constructed to include maxima, minima, ranges and lagged variables, in addition to means. The original training data set included annual data from between 1999 and 2007.

An ecosystem response model was constructed using a series of steps. Firstly, biological data were clustered to identify groups of cases (sites in a given year, referred to as a 'site-year') with similar biota. These biotic assemblages were tested for statistical distinctness, and those that were distinct were used as preliminary ecosystem states. The preliminary ecosystem states were used as the response variable for a CART analysis using the environmental data set. This identified physico-chemical associates of biota within the system. Cases within each terminal node were again tested for biological distinctness, and terminal nodes that did not support a unique biota were combined with the next closest node. The resultant CART model was then used to classify cases that had been excluded from the analysis to this point (due to the presence of missing values in the biological data set). The biota for each terminal node in the original data set were compared to the biota of the same terminal node for the new data set to ensure that the CART model was capable of predicting a biotic assemblage based on the environmental conditions for each case. The resultant ecosystem states were then characterised based on their environmental and biological characteristics.

The original annual time-step used for the analysis did not capture recent declines in condition in the Coorong sufficiently, so the model development process was repeated with more-recent data at a seasonal time step (2005-2007). The short-term (quarterly) and long-term (annual) models were then combined to give a single ecosystem state model for the Coorong. This model had eight distinct ecosystem states, with the tidal range, number of days without barrage flows, water levels, depths from the previous year and salinity as the driving variables. These states were divided into two basins; a marine basin and a hypersaline basin, each containing

four states. The threshold for the number of years without barrage flows (at 339 days) was such that it suggested a continuum of states in each basin, from healthy to degraded, and they were named as such. The eight states were named Estuarine/Marine, Marine, Unhealthy Marine, Degraded Marine, Healthy Hypersaline, Average Hypersaline, Unhealthy Hypersaline and Degraded Hypersaline.

We undertook a model evaluation process in which we predicted the ecosystem state of sites for years in the past and compared the predicted biotic assemblage to data on the actual biotic assemblage found at that location for that time. The sparseness of the data made this a difficult task, and a number of different analyses were run, firstly using data on single taxonomic groups (e.g. commercial fisheries data or bird abundances), and then using combinations of taxonomic groups (e.g. commercial fisheries data and bird abundances) to assess differences in the biotic assemblages between states and within states but between decades. The results were mixed, but suggested that, for the most part, there were distinct biotic assemblages between states, but that the biota of the 1980s was distinct from that of the 2000s. In particular, the Estuarine/Marine state appeared to have several distinct biotic assemblages, suggesting that it may, in fact, represent several distinct states over time. The model evaluation process was also undertaken for the only more-recent data set that was available; fisheries-independent fish abundances from 2008. Again, this suggested a distinct biotic assemblage in 2008 compared with those within the training data set.

A predictive state-and-transition model was developed from the ecosystem state model. By happenstance, all variables within the model were ones that could be modelled, either using catchment modelling undertaken by the Murray-Darling Basin Authority, or a hydrodynamic model developed by CSIRO. The biggest challenge involved in developing a predictive model was in extrapolating modelled data to sites that fell outside the model domain. Single and multiple regressions were used to develop these relationships. This meant that scenarios including combinations of climate change and possible management options could be developed, run through the hydrodynamic model and the ecosystem state model and then the resultant time series interrogated to predict ecosystem responses.

A second set of models was also developed for possible management actions that were designed to be an alternative to barrages flows during drought. The presence of the number of days without barrage flows as the second driving variable in both the marine and hypersaline basins meant that ecological change in response to interventions other than barrage releases (e.g. dredging) would not be adequately predicted. We, therefore, developed new models for the marine basin and for the hypersaline basin by excluding flow-related variables from the model development process. These alternative models predicted the ecosystem states of the original model with a high degree of success.

The possibility of additional ecosystem states in the Coorong was investigated. The model evaluation process suggested that the Estuarine/Marine state, in particular, may, in fact, include several distinct states. This would be consistent with the fact that the training data set included only dry years, where low flows resulted in declining ecosystem condition. Evidence suggesting the existence of healthier, more-estuarine ecosystem states was not surprising. Biological data from site-years predicted to be Estuarine/Marine revealed three or four distinct biotic assemblages. There were insufficient data, however, to resolve the environmental drivers of these assemblages, so they were not included in the predictive model. Their likely presence is, however, noted for future extensions of the model, should more data become available.

The model was subjected to a number of sensitivity analyses. The initial clustering step was one that had been identified as being key to the entire model development process, as it provided the preliminary ecosystem states. We tested the sensitivity of the model to the clustering technique used (comparing group-averaging to *k*-means clustering). We also tested the sensitivity of the ecosystem state model to random errors within the biological data set and to clustering of the environmental data set, instead of the biological data set. The clustering was more sensitive when more clusters were identified. For the long-term model, three clusters had been originally identified, and these were relatively robust, both the changes in the clustering technique and to random errors in the data set. As the number of clusters identified became greater, however, the concordance with the original clusters declined.

Overall, the model was not particularly sensitive to the choice of initial clustering technique. This was because of the independent assessment of biotic assemblages later in the model development phase identifying any groups that were not statistically distinct and allowing for these to be combined. There were mixed results when clusters based on the environmental data set were compared with those based on biological data, with the short-term data set showing good concordance and the long-term data set having relatively poor concordance, despite a small number of clusters. The decision to use the biological data set was based on the modelling and monitoring capabilities of management agencies and the notion that biotic assemblages usually respond to physico-chemical parameters, rather than the other way around.

We also assessed the sensitivity of the predicted ecosystem states to the exact value of the threshold for each driving variable. It was not surprising that predictions were most sensitive to the value of thresholds for variables closest to the top of the decision tree (i.e. tidal range and number of days without flow). The model was very robust to changes in the threshold value for variables near the bottom of the decision tree (i.e. water levels and salinity). However, none of the threshold values were statistically significant when compared with a distribution built from random choices of the threshold value.

The model was used to assess ecological response to a series of scenarios. These scenarios included testing the effects of climate change, current extraction levels, sea level rise, The Living Murray initiative, dredging at the Murray Mouth, and increased flows through Salt Creek via the Upper South East Drainage scheme. Methods are presented here, but all results and interpretations based on these scenarios can be found in Lester *et al.* (2009).

The ecosystem state model developed here has a number of limitations. These include our inability (as yet) to characterise additional ecosystem states for truly estuarine conditions or to build spatiotemporal links into the predictive model. The model was built using a sequence of low flow years, so it is biased towards predicting the decline of the Coorong ecosystem rather than any recovery. This means that potential recovery trajectories and any lags may not be adequately described by the model as is. Also it is not intended to describe causal links, but rather robust correlations between the environmental and biological aspects of the ecosystem for use as indicators by managers.

Nonetheless, the ecosystem state model seems to be a useful tool in predicting ecosystem response, and in understanding some of the physico-chemical drivers of biotic assemblages. It has the potential to improve the management of large ecosystems, by dramatically simplifying definitions of ecosystem health and allowing management for a set of ecosystem states, rather than arbitrarily-defined thresholds. The range of environmental conditions over which given ecosystem states occur could replace attempts to define limits of acceptable change based on the literature or untested expert opinion. The ecosystem state model is a data-driven, multivariate representation of a complex system that can be used in simple assessments of ecological condition, or in detailed scenario analyses.

1.   Introduction

Ecosystem response modelling is a relatively new tool in the natural resource management (Sutherland, 2006). With streams of environmental data becoming increasingly automated, more data are now being collected than ever before. With improved computer-processing power, new techniques in ecosystem modelling continue to emerge, taking advantage of this increased computational capacity. This means that, for a given problem, there are likely to be a range of ecological modelling options available. A comprehensive review of the approaches commonly used to predict the ecological consequences of environmental change was undertaken by Sutherland (2006). He recommended a combination of traditional model development to address a specific problem with the use of expert opinion, output-driven modelling and scenario analyses (Sutherland, 2006). This recommendation presents ecologists and modellers with a challenge to develop new techniques and/or apply several of them together in novel combinations. Here we attempt to answer this challenge for the Coorong estuary and lagoons.

The Murray-Darling Basin ends on the South Australian coast in an estuarine and coastal lagoon complex called the Coorong (Figure 1.1). The Coorong is a long, narrow system comprising two main lagoons, the North Lagoon and the South Lagoon, and an estuarine region around the Murray Mouth, which is the single connection to the Southern Ocean (Phillips and Muller, 2006). The system is an inverse estuary (Wolanski, 1987), with both the Murray Mouth and the source of freshwater, the River Murray, at the same end, rather than the more common pattern where freshwater enters at the upstream end and flows to a downstream mouth. The water sources into the Coorong include freshwater flows from the River Murray across a series of barrages, marine water as tidal input through the Murray Mouth, fresh to brackish inputs from the Upper South East Drainage (USED) scheme of southeastern region of South Australia via Salt Creek near the southeastern end of the system, and more localised rainfall and groundwater inputs (Webster, 2006).

The Coorong is ecologically significant as a part of the Coorong, Lower Lakes and Murray Mouth (CLLAMM) region that is listed under the Ramsar Convention on Wetlands of International Importance (Department for Environment and Heritage, 2000) and is designated as one of six icon sites within the Murray-Darling Basin (Phillips and Muller, 2006). For example, the region supports more than 1% of the global population (one of the criteria for Ramsar listing) for a number of migratory wading bird species, including the curlew sandpiper (*Calidris ferruginea*), red-necked stint (*Calidris ruficollis*) and sharp-tailed sandpiper (*Calidris acuminata*) (Paton *et al.*, in press). The region also has significant cultural, economic and recreational values, with sizable local tourism and commercial fishing industries, nearby agriculture and a local indigenous Australian community, the Ngarrindjeri nation.

There are a number of potential 'levers' able to be used in the management of the system to affect ecosystem condition. These levers include: environmental flows of fresh water, either from the USED via Salt Creek, or from the River Murray; dredging of the Murray Mouth to maintain the connection to the ocean; barrage operations controlling the flow of water between the Lower Lakes (Figure 1.1) and the Coorong, including the operation of fishways; and various local engineering solutions that are possible in the region, including pumping and dredging options. These, combined with climatic forcing factors, determine the ecological condition of the Coorong, and the identified "water benefits" of the region, such as a viable fishery based on fish recruitment, tourism, cultural values, sustained waterbird populations and an open Murray Mouth.
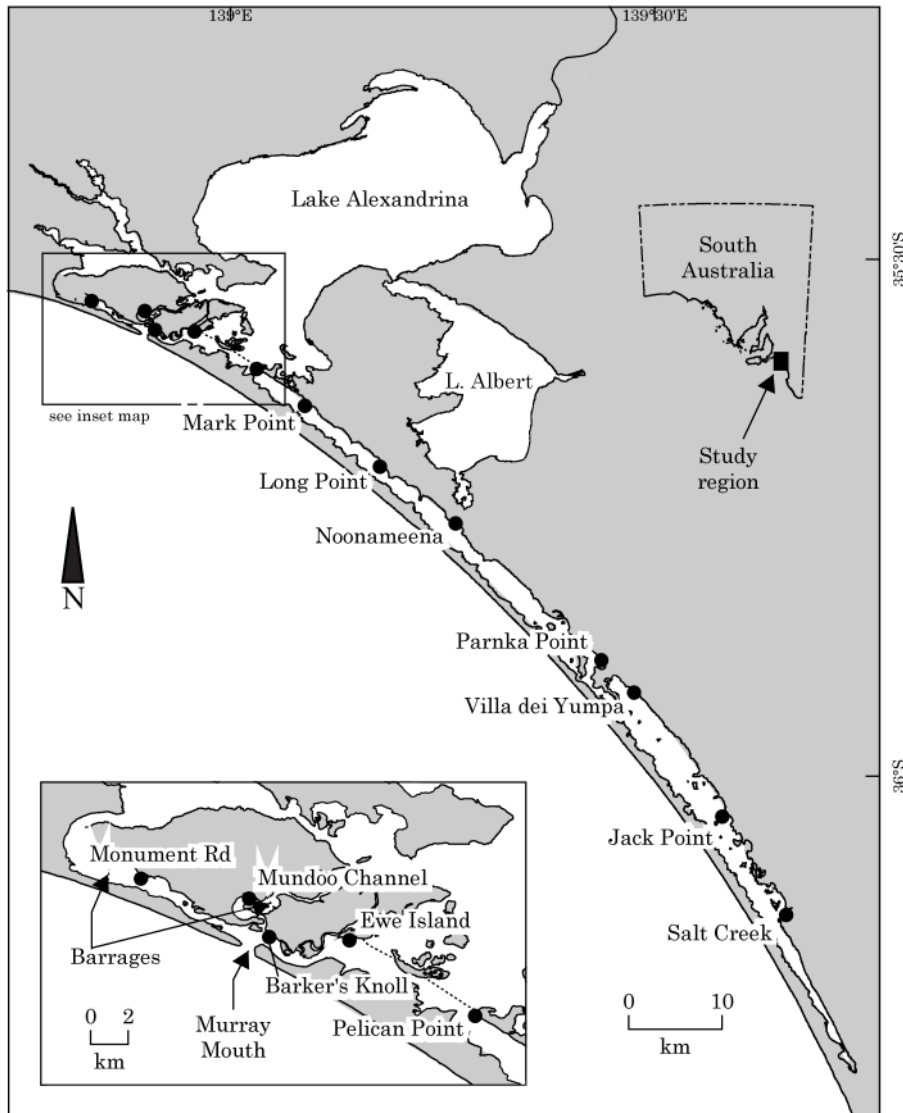
Figure 1.1. Map of the Coorong showing the twelve study sites used as focal locations during CLLAMMecology and forming the basis of our ecosystem response modelling

(Source: Craig Noell, SARDI Aquatic Sciences, South Australia)

The recent observed decline in condition (Lester and Fairweather, in press; Paton *et al.*, in press) and the desire to have a good scientific basis to guide the management of the system prompted the formation of the CLLAMMecology Research Cluster (Lamontagne *et al.*, 2004). CLLAMM is an acronym for 'Coorong, Lower Lakes and Murray Mouth', thereby describing the overall region in which the Cluster was to operate. The Cluster included researchers from the University of Adelaide, Flinders University, the South Australian Research & Development Institute Aquatic Sciences (SARDI Aquatic Sciences) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Water for a Healthy Country Flagship. Management agencies responsible for the Coorong were also involved, including the South Australian Department for Environment and Heritage, and the South Australian Department of Water, Land and Biodiversity Conservation. The aim of the Research Cluster was to develop an ecosystem-level understanding of the Coorong, Lower Lakes and Murray Mouth. Four themes were developed to achieve this aim, targeting the response of key species, productivity and trophodynamics in the system, dynamic habitat availability, and

ecological responses to a range of alternative futures. This final theme, labelled CLLAMM Futures, aimed to combine existing information with that derived from during CLLAMMecology to develop an ecosystem response model for the Coorong.

The CLLAMM Futures theme thus aimed to develop a set of models at an ecosystem scale to integrate the knowledge generated by the other CLLAMMecology themes with existing knowledge. These models developed are system-wide models that have been used to investigate the likelihood of a number of future scenarios, including possible management options and alternative states within the Coorong region (Lester and Fairweather, 2009a). The models are explicit in both space and time and allow prediction of the response of the ecosystem to the identified system drivers (including climatic forcing and management options, for example). This report describes in detail the methods used to develop and then apply these models to a set of scenarios of interest.

This report is one of a series of outputs from the CLLAMM Futures theme. A "road-map" outlining the other outputs from the theme (either already produced or in preparation) can be found in Appendix A.

2.   Methods

2.1.   Method development

2.1.1.   Lessons from the modelling review

For a complex modelling application, such as that undertaken CLLAMM Futures, we considered that the best approach was likely to include a combination of several different modelling techniques. Achieving the objective of the CLLAMM Futures theme, to construct a single, system-wide response model for the region, involved meeting a number of considerable challenges. The data sets that were available for use in modelling were disparate in nature and collection methods, included qualitative, semi-quantitative and quantitative data, were patchy in both space and time, but did not cover all variables of interest. The nature of the data, along with the complexity of the system as a whole, with estuarine, marine and hypersaline regions, meant that constructing a valid deterministic model was not feasible. There was simply insufficient information for any deterministic model to adequately capture the relevant processes and to be properly validated. For this reason, the CLLAMM Futures theme focused on developing a statistical, stochastic model, based on probabilities and analyses of existing data sets, rather than attempting to mathematically characterise all the interacting processes occurring.

We undertook a review of the modelling and analysis alternatives considered most promising for use during CLLAMM Futures (Lester and Fairweather, 2008b). Based on this review, a number of analytical techniques appeared promising for use during CLLAMM Futures. A summary table listing the advantages, disadvantages and potential applicability to CLLAMM Futures of the 13 possibilities assessed is presented in Appendix B. Classification and regression tree (CART) analysis (where the variation in a univariate response variable is explained through repeated splitting of the data into increasingly homogenous groups based on a variety of exploratory variables; De'ath and Fabricius (2000), multiple regression tree (MRT) analysis (an extension of the univariate regression trees used in applications such as CART, where the univariate dependent variable is, in effect, replaced by a multivariate response; De'ath, 2002), and structural equation modelling (SEM) (an extension of path analysis that seeks to minimise the differences between the variance and covariance matrices of the model under investigation; Reckhow *et al.* 2005) were identified as good initial choices, given their ability to classify and rank independent variables. This will allow the various states postulated for the CLLAMM region to be objectively defined and tested. Of the three methods initially identified, only CART analyses were finally incorporated into the ecosystem state model described here due to the intensive data requirements of SEM and difficulties in the identification of groups in the biological data (especially across multiple data sets) using MRT.

The information from these analyses was then used to construct a state-and-transition (S&T) model describing the ecosystem states of the CLLAMM region. S&T models draw on the theory of alternative states for a community (Westoby *et al.*, 1989; Bestelmeyer *et al.*, 2004; Wilkinson *et al.*, 2005), although the states do not necessarily need to be true alternative stable states. The technique uses a box-and-arrow flow diagram to represent an ecosystem with the boxes used to define discrete system states, and the arrows representing the events and processes that cause the system to move between these states (transitions) (Plant and Vayssieres, 2000). S&T models are designed to be a framework for allocating a system to various states.

Bayesian belief networks (BBNs) were identified as one possibility for constructing the S&T model, given that it was likely to include information from many different analyses, of varying

degrees of reliability, including data from experiments, as well as from condition monitoring, along with information from the literature and expert opinion, anecdotal information and suspicions where data were unavailable. In the end, this was not used, with a more traditional S&T structure being employed, in an attempt to maintain the data-driven integrity of the model (i.e. all states and transitions are derived from the available data), but this approach still has possibilities for future attempts at similar modelling applications.

A diagram illustrating how the various modelling techniques were combined in the ecosystem state model is shown in Figure 2.1. This synthesis of modelling techniques was then used as the basis of scenario modelling (Lester *et al.*, 2009). The process of converting the ecosystem state model into a predictive model and running scenario analyses is described below.
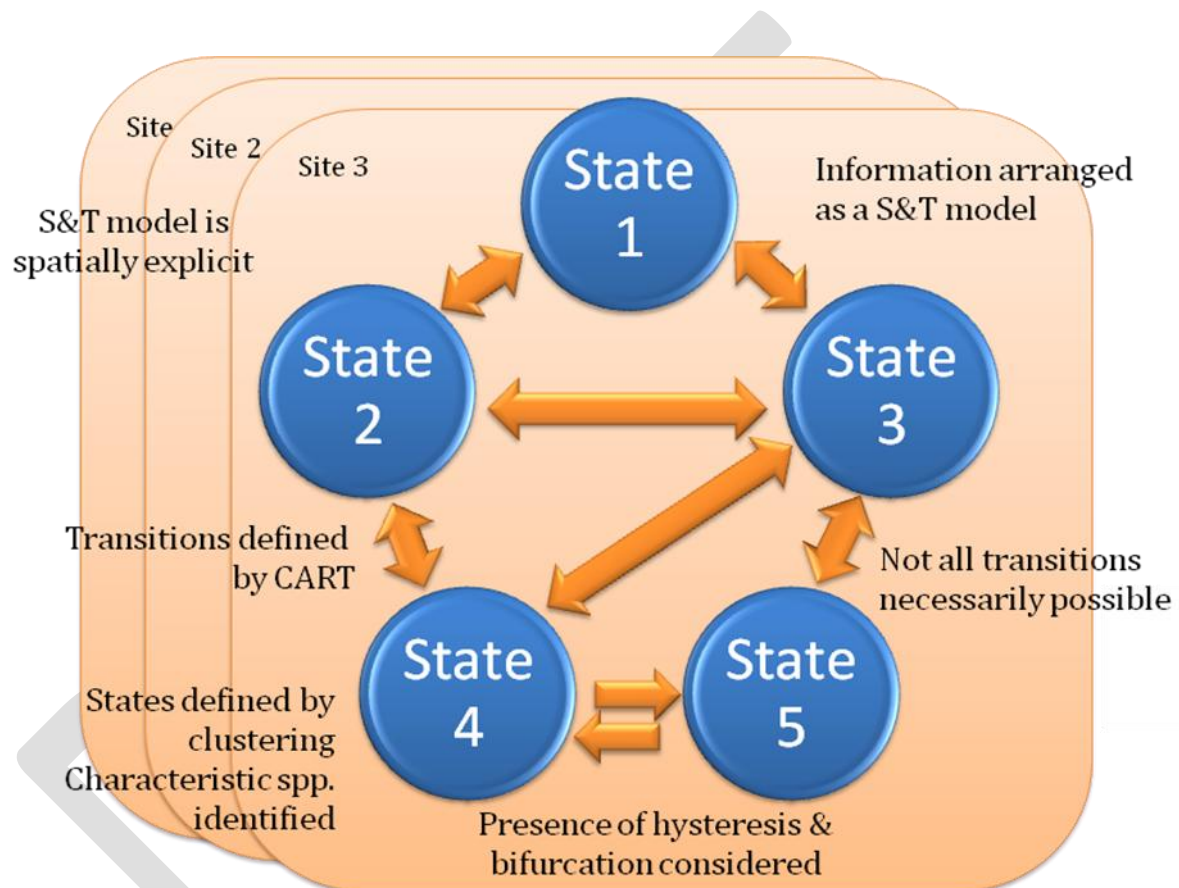
Figure 2.1. Synthesis of modelling techniques used to construct the ecosystem state model

Note that this figure shows a hypothetical S&T model. Not all transitions between states are possible. Some states show hysteresis, where the driving variable (or threshold) governing transitions in one direction is different from that in the other direction. The model is spatially explicit (indicated by the different squares for each site).

2.1.2.   Data collection

As many data as possible were compiled for the Coorong. Data sources included the available literature, data collected by CLLAMMecology researchers either prior to or during CLLAMMecology, monitoring data that had been collected by state or federal agencies in the region, and outputs of other models constructed for the Coorong.

We divided the data into a biological data set and an environmental (physico-chemical) data set. The biological data set consisted of abundance data for birds, fish and benthic macroinvertebrates, plant coverage and abundance of propagules for the dominant macrophyte species, *Ruppia tuberosa*, and catch per unit effort data for the local commercial fishery (Table 2.1). The environmental data set included modelled water levels, depths and salinities along the length of the Coorong, daily tidal range, flows over the barrages, meteorological data, and measured water quality parameters, including the concentration of nutrients, pH, alkalinity and turbidity (Table 2.2). Tables 2.1 and 2.2 give references relating to the data-collection methods and attribute the data to their various provenances. Due to these varied provenances, we have reported values in their original units rather than attempting to convert them where more standard values are common or more recent.

For the environmental data set, a range of parameters related to variability *per se* were calculated for each variable, including maxima, minima, ranges, variability, moving averages (3 years) and lagged effects (1 and 2 years), particularly for flow and depth characteristics. This variety of parameters was included based on the understanding now that the mean conditions are not always the best predictors of ecosystem state or condition (Gaines and Denny, 1993).

Category	Metrics included	Variable	Units	Provenance	Notes
Macrophytes	Percent of cores containing *Ruppia*, number of shoots, number of type I turions, number of type II turions	*Ruppia* distribution	%, number per core	University of Adelaide	See Paton and Rogers (2008) for a description of sampling methods. Included in long-term analyses only. A single survey was undertaken each year
Bird abundances	Average number sighted	Species	abundance per site	University of Adelaide, DEH	See Paton *et al.* (in press) and Dittmann *et al.* (2006) for a description of sampling methods. Species where >500 individuals were sighted over 8 years were included in long-term data, and >100 individuals sighted over the 3 years for the short term
Fish abundances	Average CPUE	Commercial fish catch	CPUE (net days) per site	PIRSA	The commercial fish CPUE is based on data collected from commercial fishermen on a daily basis. Only species with a total of >5 CPUE across all sites and times were included. Averages were calculated of annual CPUE for long-term analyses and quarterly CPUE in the short-term
	Average abundance	Small-bodied fish abundances	abundance per site	SARDI Aquatic Sciences	Included in short-term analyses only. See Noell *et al.* (2009) for a description of sampling methods.
Benthic macroinvertebrate abundances	Average number of individuals per species	Adult abundances	abundance per site	Flinders University	See Dittmann *et al.* (2008) for a description of sampling methods. Included in short-term analyses only
	Average number of individuals per species	Juvenile abundances	abundance per site	Flinders University	Included in short-term analyses only

Table 2.1. Biological variables included as part of the input data sets and their provenance

Note: All variables were included in both long- (annual time-step) and short-term (quarterly time-step) analyses unless specified. Where more than one survey was undertaken within the time period analysed, abundance (or other relevant variable) were averaged. The Units column presents the numerical units used in the analysis, and gives the sampling units in which the data was collected (e.g. per site for invertebrate abundances). The division of turions into types I and II follows (Paton and Rogers, 2008). Catch per unit effort was calculated using the number of nets by the number of days each was set (net days) as the measure of effort.

Category	Metrics included	Variables	Units	Provenance	Notes
Meteorology	Average minimum, average maximum, maximum, minimum	Temperature	^{o}C per day	BOM	Data from two weather stations (51 km apart) were used for all meteorological variables: Hindmarsh Island & Meningie, SA. Averages were taken of daily measurements
	Average, maximum & minimum	Relative humidity	% twice per day		Measured at 9 am & at 3 pm daily. Averages were of daily values for each reading
	Average, maximum daily, minimum daily, total across the time period	Precipitation	mm per day		Averages were of total daily precipitation
	Average, maximum daily & minimum daily. Average, maximum & minimum gust wind speed at Hindmarsh Island	Wind speed	knots twice per day		Wind speed measured at 9 am & 3 pm daily. Gust wind speed is the maximum gust recorded each day. Averages were of daily values for each reading
	Average, maximum, minimum	Cloud coverage	eighths per day		Averages were of average daily cloud coverage
Water quantity	Average, average from previous year, average from 2 years ago, 3-year moving average, average minimum, average maximum, maximum, minimum, variance, change in water level across the year	Water depth	m per hour	CSIRO	Modelled data produced from a hydrodynamic model described in Webster (2006). Averages were of average daily water depths
	Average, average from the previous year, average from 2 years ago, 3-year moving average, average minimum, average maximum, variance, minimum & maximum	Water level	m AHD per hour		Modelled data produced from a hydrodynamic model described in Webster (2006). Averages were of average daily water levels

Table 2.2. Environmental variables included as part of the input data sets and their provenance

Note: All variables were included in both long- (annual time-step) and short-term (quarterly time-step) analyses unless specified. Where more than one survey was undertaken within the time period analysed, abundance (or other relevant variable) were averaged. The Units column presents the numerical units used in the analysis, and gives the sampling units in which the data was collected (e.g. per day for rainfall)

Category	Metrics included	Variables	Units	Provenance	Notes
	Average, maximum	Daily tidal range	m per hour		Modelled data produced from a hydrodynamic model described in Webster (2006). Averages were of daily tidal range
	Average daily flow, average from the previous year, number of days in the time period when flow occurred, maximum number of days since flow, average flow size per event, average flow size from the previous year, average total flow & total flow from the previous year	Flow	GL per day	MDBC	Modelled data produced by MSM-BigMod (Close and Sharma, 2005)
Water quality	Average, average maximum, average minimum, variance, average from the previous year, average from 2 years ago, 3-year moving average, average & maximum change in salinity across the year	Salinity	$g\ L^{-1}$ per site	CSIRO	Modelled data produced from a hydrodynamic model described in Webster (2006). Averages were of average daily salinities
	Water temperature where and where invertebrates were collected	Temperature	^{o}C per site	Flinders University	Measured data, only included for short-term analyses
	Average, maximum & minimum total phosphate, total Kjeldahl nitrogen & ammonia concentrations	Nutrient concentration	$mg\ L^{-1}$ per site	DEH	Data for all remaining water quality parameters were measured quarterly. Averages were of quarterly data for long-term analyses. For short-term analyses, a single measurement was included

Table 2.2 cont. Environmental variables included as part of the input data sets and their provenance

Note: All variables were included in both long- (annual time-step) and short-term (quarterly time-step) analyses unless specified. Where more than one survey was undertaken within the time period analysed, abundance (or other relevant variable) were averaged. The Units column presents the numerical units used in the analysis, and gives the sampling units in which the data was collected (e.g. per day for rainfall)

Category	Metrics included	Variables	Units	Provenance	Notes
	Average, maximum, minimum	pH	pH units per site		
	Average, maximum, minimum	Alkalinity	mg L^{-1} per site		
	Average, maximum, minimum	Turbidity	NTU per site		
	Average, maximum, minimum	Electrical conductivity	mg L^{-1} per site		
	Average, maximum, minimum	Conductivity	µS cm^{-1} per site		
	Average, maximum, minimum	Chlorophyll a & b concentration	µg L^{-1} per site		

Table 2.2 cont. Environmental variables included as part of the input data sets and their provenance

Note: All variables were included in both long- (annual time-step) and short-term (quarterly time-step) analyses unless specified. Where more than one survey was undertaken within the time period analysed, abundance (or other relevant variable) were averaged. The Units column presents the numerical units used in the analysis, and gives the sampling units in which the data was collected (e.g. per day for rainfall)

2.1.3. Developing the basic ecological response model

The initial ecological response model was constructed based on an annual time-step at each of the 12 focal sites identified for CLLAMMecology (Figure 1.1). Data were aliased where required to maximise the spatial and temporal coverage of the data sets to the relevant scales. This involved aligning data from the various regions and sites used during the collection of each data set with the focal sites identified for CLLAMMecology, and averaging data where multiple sampling events occurred within a single time-step (e.g. per annum). Maximum data was available for the period 1999 to 2007, which were the years included in the original model development. This model is referred to as the 'long-term' ecosystem state model. This initial long-term model did not adequately capture declines in ecosystem condition that were identified during CLLAMMecology (Lester and Fairweather, in press; Noell *et al.*, 2009; Paton *et al.*, in press), so a second model was constructed using a quarterly time-step for the years 2005 to 2007, where the decline in condition of the Coorong ecosystems was most-clearly evident (referred to as the 'short-term' ecosystem state model). These models were eventually combined to produce a single ecosystem state model for the Coorong as described below.

Developing an ecological response model (ERM) for the Coorong involved a number of steps. The model development process is summarised in Figure 2.2. Each step is described in detail below.

Identifying preliminary states (Steps 1 & 2)

Cluster analysis was used to identify preliminary groups of cases that had similar biotic assemblages. Only cases that included a complete set of biological data were able to be clustered. Step 1 (Figure 2.2) involved splitting the data into those cases that were complete ($n = 85$) and could be clustered (Step 2) and those that had missing values ($n = 23$) and would be retained for model validation later (Step 5).

For the long-term time analysis, the data set included macrophyte cover, bird abundances and CPUE for commercial fish species. The cluster analysis (Step 2) used a group-average linkage algorithm and a SIMPROF test to identify clusters which were statistically distinct (Clarke and Gorley, 2006). These analyses were undertaken in PRIMER 6. This analysis identified four distinct clusters within the long-term data, one of which consisted of only a single case. The division between this case and the rest of the data was the first division of the data, so that case (at Barkers Knoll in 2001) was excluded from further analyses as an outlier, and so three clusters were retained.

The short-term data set included bird abundances, CPUE for commercially-fished species, fisheries-independent fish abundances, and abundances of juvenile invertebrates, all at a quarterly time step ($n = 144$; i.e. 12 sites by 3 years by 4 quarters). Migratory birds were excluded from the data set because their presence is seasonal and could be largely unrelated to the condition of the Coorong. There was very little overlap in the cases that included data for both fisheries-independent fish abundances ($n = 49$) and juvenile invertebrate abundances ($n = 69$), so we divided the data into two (excluding cases that still had missing values; Step 1) and ran cluster analyses on both halves (Step 2) in PRIMER 6.0. Group-average cluster analysis on the short-term data including juvenile invertebrate abundances identified six initial clusters, two of which consisted of only a single case each. As for the long-term analyses, these two cases were excluded from further analysis, leaving four statistically-distinct biological clusters for the next step in the model development process.

The short-term data set including fisheries-independent fish abundances proved to have nine distinct biotic groups. Of these, four consisted of either a single case, or two cases, none of which was sufficient to be subjected to a CART analysis (Step 3), so these seven
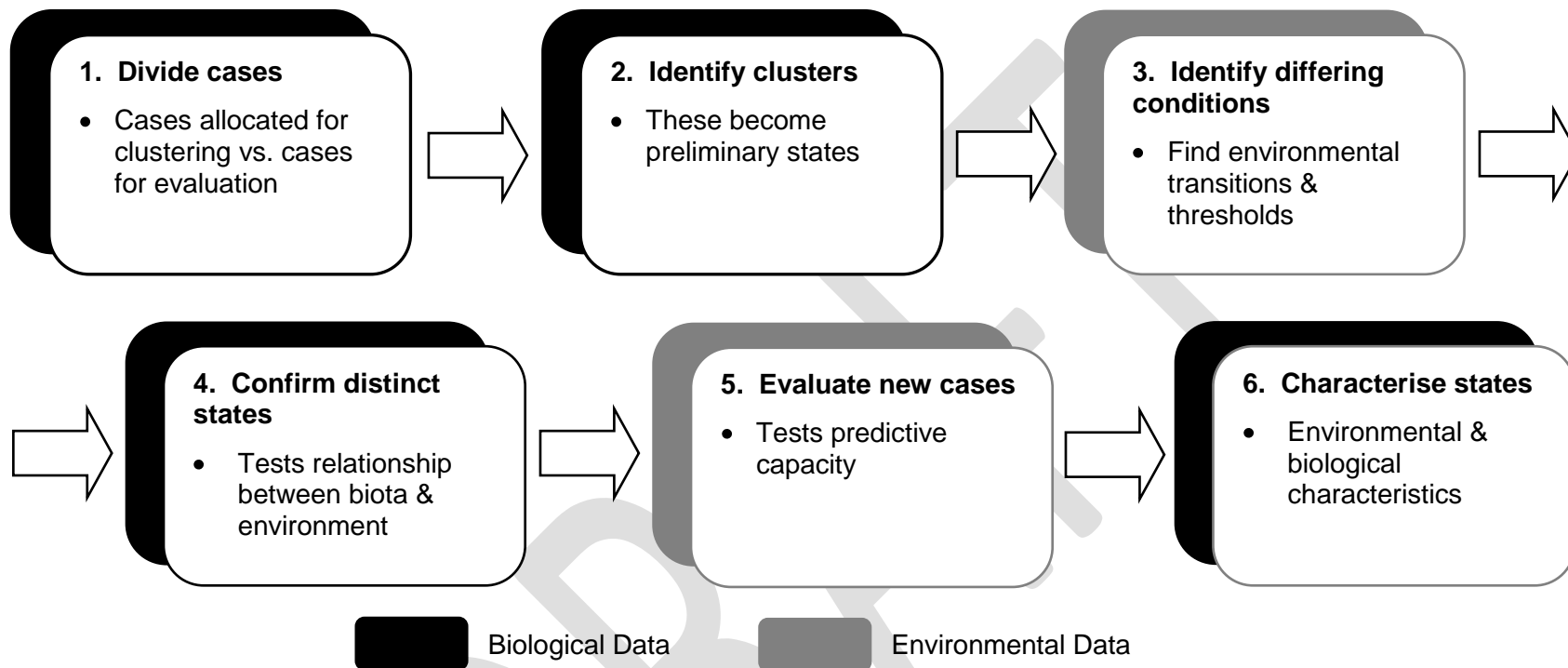
Figure 2.2. Flow diagram summarising the steps in developing an ecosystem response model for the Coorong using both environmental and biological data sets (as per Tables 2.1 & 2.2)

Note: This process was followed for both the long-term (annual time-step) and short-term (quarterly time-step) model development. These were then combined into a single ecosystem model for the Coorong.

cases were excluded from further analyses. The cases that were excluded were from Quarters 2 and 3 in 2007 at Monument Road, Quarter 3 in 2007 at Barkers Knoll and Ewe Island, Quarter 1 from 2007 at Jack Point and Quarters 1 and 2 in 2007 from Salt Creek. This left five distinct biological clusters for the next step in the model development process.

Before proceeding to the next step, the biological clusters identified in the two subsets of the short-term data set were combined. Where a single case was categorised in both the short-term cluster analyses (i.e. had abundance data for both fisheries-independent fish and juvenile invertebrates), the two clusters to which it belonged were amalgamated. This gave us a total of six identified clusters across the two analyses for the short-term data set. This approach may have resulted in more clusters being identified than would have been supported using a single complete data set, but later steps in the process were designed to identify extraneous states and combine these where needed (Step 4), so this was not seen to be a problem at this stage of the process.

These biologically-distinct clusters of cases that were identified were considered to be preliminary states for the purposes of moving to Step 3.

In addition to group-average cluster analysis, a number of other analytical techniques were explored to identify groups of biota that may constitute preliminary states within the model. Multiple regression tree analyses were attempted in order to identify biological groups and indentify the environmental variables driving these distinct assemblages simultaneously. These analyses were undertaken using the 'mvpart' library in the R statistical environment (De'ath, 2002). This attempt was abandoned when the analysis was not able to detect any subsets within the biological communities, despite the differences found using cluster analysis. This may have been due to the large number of biological variables, and the degree of non-normality within the data set. Principal components analysis was also explored as another tool for identifying distinct biological assemblages within the long-term data set. These analyses were undertaken using SYSTAT 11. The various principal components were not found to have strong links to the environmental variables when used during CART analysis (see below for detailed methods), so this avenue was not pursued further.

Other clustering techniques were also explored. *K*-means clustering was not considered useful in the first instance, because it required an *a priori* selection of the number of clusters to be identified. Given the exploratory nature of this part of the analysis, we preferred to allow the analysis to identify the number of significant clusters, rather than imposing any preconceived ideas of the ecosystem state diversity of the system on the model. Complete-linkage and single-linkage clustering were also explored. Single-linkage clustering tended to identify clusters consisting of a single or very small number of cases, which was not conducive to identifying broad environmental parameters that drove differences in biotic assemblages. Complete-linkage clustering generated a more even distribution of cases across clusters and was pursued through to the investigation of environmental drivers, but subsequent CART analyses did not produce models with the same goodness of fit as for the group-average-derived preliminary states. The complete-linkage clustering option does merit consideration in future attempts to apply this modelling technique, however, to gain an understanding of whether it out-performs group-averaging under any other circumstances. These cluster analyses were all undertaken using PRIMER 6.

Differentiating between preliminary states using environmental variables (Steps 3, 4 & 5)

Classification and regression tree analyses were used to differentiate between the preliminary states using the environmental conditions under which each state occurred (Step 3, Figure 2.2). These analyses were undertaken using CART 6.0 (Steinberg and Golovnya, 2007). The package 'rpart' in the R statistical environment (Therneau and Atkinson, 2009)

was also used to run CART analyses, but difficulties in penalising potential explanatory variables for the frequency of missing values meant that this platform was abandoned in favour of using CART 6.0.

In the first instance, the predictive capacity of the various properties of water quality and meteorological variables was tested (by running multiple CART analyses in CART 6.0). The data set included means, minima and maxima for all variables. We did not believe that it would be useful to include all variables as possible predictive variables, because the likelihood of correlations between the mean, minimum and maximum of a single variable was high (as was indicated by preliminary analyses). By using each set of variables in turn, we established that the maxima had the highest predictive capacity, so these variables were included, while the minima and means were omitted. For flow, water quality and salinity variables, a more complex set of variables was included (including lagged variables, variances and flow within particular sets of months). This made it difficult to identify in advance which variables were likely to be inter-correlated, so correlations amongst predictive variables were undertaken after models had been constructed. Where similar variables were measured by more than one technique (e.g. conductivity and electrical conductivity (EC) were both measured and salinity was modelled), the source with the most-complete data was used and the others excluded (i.e. modelled salinity was included because conductivity and EC were measured quarterly at best). Where two significant predictor variables that were identified by the CART model were significantly correlated, the variable that explained the smaller proportion of the variance was excluded and the model re-run, until no predictor variables retained within a model were significantly correlated.

Twoing splitting was used during the CART analysis because it is the most likely criterion for creating child nodes of relatively equal sizes (Steinberg and Golovnya, 2007). The best tree was selected using the One Standard Error (1SE) rule (Breiman *et al.*, 1984), and only cases for which a preliminary state was identified were included in the analysis. Cross-validation occurred with seven folds, as this was the size of the cluster with the smallest number of cases. The minimum number of cases allowed was set to five for a parent node and two for a child node. To penalise variables with missing values, β was set to 0.6 (Steinberg and Golovnya, 2007).

For the long-term data set ($n = 85$), a three-node tree was produced at a relative cost of 0.348 (with lower relative costs being more favourable; Steinberg and Golovnya, 2007). The two splitting variables were the average annual modelled salinity and the average annual depth from the previous year. This model correctly classified 95% of the cases in the learning data file, and 86% of the cross-validated cases. ANOSIM analyses in PRIMER 6.0 of the biotic data set confirmed that the cases in the three terminal nodes were, in fact, biologically distinct (Step 4; Global $R = 0.796$, $p = 0.001$; all pair-wise comparisons had $p = 0.001$).

Cases that had been excluded from the cluster analysis step due to some missing data (at Step 1; $n = 23$) were used to validate the predictive capacity of the CART model (Step 5). The terminal node for each of these cases was predicted using the CART decision tree. For each terminal node, the biological assemblage for the original complete cases used in Step 3 was compared with these new cases that had previously been excluded from the model development process (at Step 1).

In order to compare the biological assemblages, two nested subsets of the long-term biological data were created. This involved separating the CPUE data for the commercial fishery (which included all 108 cases) and then the CPUE data with the bird abundance data (which included 96 of 108 cases). The data relating to *Ruppia* cover had the fewest cases, so they were excluded from this step. ANOSIM analyses (in PRIMER 6.0) were used on the terminal node membership of each case, with pair-wise comparisons indicating whether the biota of cases from the original terminal node were statistically different from that of the new cases. Identifying a significant difference in one of the two ANOSIM analyses (on the two

data sets) was considered sufficient to confirm a distinct biotic community. No significant differences were identified between new and original cases for any of the three terminal nodes in either analysis. The long-term CART decision tree was therefore considered sufficiently predictive with respect to predicting the biotic community of a site-year based on its environmental conditions.

As touched upon earlier, this model adequately described the spatial distribution of communities in the Coorong over the nine-year period, but it did not show the decline in condition that was observed by CLLAMMecology researchers in the latter years (2005 onwards; Lester and Fairweather, in press; Noell *et al.*, 2009). In order to adequately assess the ecological condition of the Coorong, and the likely response to management scenarios, it was important that we could accurately predict abrupt temporal changes in the ecosystem function. For this reason, the short-term analysis was also conducted.

Using the combined group-average cluster results as the preliminary states with the short-term biological data set (as described above), we repeated the CART analyses to identify the environmental variables driving these preliminary states in CART 6.0 (Step 3).

The same predictive variables were used for the short-term CART analysis as for the long-term analysis (although the values were re-calculated for the quarterly time-step). Twoing splitting was again used, with the 1SE rule used to select the best tree. Cross-validation occurred with four folds, as this was the size of the cluster with the smallest number of cases. The same parameters governed the minimum size of parent and child nodes and the penalty for missing values.

CART identified a tree with nine terminal nodes and a relative cost of 0.428. The splitting variables were the average daily range in modelled water levels (mostly due to tides), the maximum number of days without flow over the barrages, the average modelled water level, the average annual modelled salinity from the previous year, the average modelled water depth from the previous year and finally the average modelled water level from the previous year. This model correctly predicted 87% of cases in the learning data set and 66% of cases in the cross-validation.

The patchiness of the biological data set made testing the distinctness of these eight terminal nodes somewhat complex (Step 4). The short-term data set was divided into a number of (mostly) nested subsets (i.e. including CPUE from commercially-fished species, CPUE and bird abundances, CPUE and bird and juvenile invertebrate abundances and finally CPUE and bird and fisheries-independent fish abundances). The first subset (CPUE only) included all 144 cases, but of course had a very limited subset of the available biological data. The final two subsets (CPUE with bird and juvenile invertebrate abundances ($n = 62$) and CPUE with bird and fisheries-independent fish abundances ($n = 49$)) had the greatest range of available biological data, but smaller numbers of cases each. ANOSIM-analyses were undertaken in PRIMER 6.0 on each subset in turn. Where evidence existed that the cases grouped into a terminal node did constitute a biologically-distinct community (by having a significant difference in a pair-wise comparison, with some consideration given to small samples sizes), those terminal nodes were considered distinct. Where no ANOSIM analysis indicated that terminal nodes were significantly different, they were combined.

This process resulted in the combination of two terminal nodes from the short-term model, leaving a model including eight terminal nodes. The terminal nodes combined represented the last split on the existing short-term classification tree (with a splitting variable of the average modelled water depth from the previous year), so the existing short-term tree could be pruned, and constructing a new tree was not necessary. The pruned tree correctly predicted the preliminary state of 79% of the training data set and 64% of the test data set under cross-validation.

As for the long-term model, the cases excluded at Step 1 ($n = 48$) were used to assess the ability of the CART model to predict the biological community of a site-season, using

ANOSIM analyses in PRIMER 6. The short-term CART model (in CART 6.0) was used to predict terminal nodes for each of the new cases based on its environmental characteristics. The various subsets of the short-term biological data described in Step 4 (with the exception of the two including the juvenile invertebrate and fisheries-independent fish abundances) were used to compare the biotic assemblages present in the new set of cases for each terminal node with the original cases used to define that node. All terminal nodes showed non-significant differences between the biota of the new and old cases except for one. This node was not the final split in the CART model (as occurred during Step 4), so was not easy to identify an additional splitting variable to separate the new cases from the original cases. There were also relatively few cases upon which to base such a decision (56 possible permutations in the ANOSIM analysis). At this point in the model development process, the node was left as it stood, with a flag that it may in fact represent more than one ecosystem state in the system.

Combining the long- and short-term models

Developing a single ecosystem model for the Coorong was the objective of the CLLAMM Futures project, so the long-term and short-term models were combined into a single model spanning both time-steps.

ANOSIM analyses were undertaken in PRIMER 6.0 on all cases from both the data-sets (long-term and the short-term) to identify the states that were biologically distinct. This necessarily involved including some data more than once within the analyses (as 2005 to 2007 occurred in both data sets), although the time-steps analysed, and hence the averages, were different. However, the goal of the exercise was to combine states that occurred in both models, so these repeated data were less likely than independent data to be in significantly-different states across the two models.

Catch per unit effort data for the commercial fishery was the only data set that was collected in a consistent manner across both the long-term and short-term data sets. CPUE is also a method that is standardised for differences in effort levels, so it adjusted for the difference in time-step between the two data sets. An ANOSIM analysis was undertaken on all cases from both the data sets, assessing pair-wise differences between the terminal nodes identified in Step 3 (and confirmed through Steps 4 and 5, Figure 2.2) for each of the long-term and short-term models. This analysis indicated that the biotic assemblages for cases in two of the three terminal nodes from the long-term model were not significantly different from the biota of the cases from a number of the terminal nodes in the short-term model. This suggested that the variation in the long-term model for those terminal nodes was partitioned across several short-term terminal nodes (which in fact, represented variation that was the rationale for constructing the short-term model in the first instance). One of the long-term terminal nodes, however, had cases where the biotic assemblage was significantly different from all cases associated with short-term nodes.

Cases from the distinct long-term node were combined in an environmental data set with all of the short-term cases ($n = 213$). A number of variables were excluded from the data set, because the difference in time-step affected their relative values. These included the total precipitation over the time-step and a number of flow variables that were based on total flow over the time-step.

Step 3 was repeated using this combined long- and short-term environmental data set to develop a combined CART model in CART 6.0. The same set of options governed the choice of tree, cross-validation and penalty for the proportion of missing values in predictive variables. The response variable to be predicted was the terminal node identified in either the long-term or the short-term CART model (whichever was relevant). Where significant predictive variables were correlated with each other, the variable explaining a lower proportion of the variance in the model was excluded, and the model was re-run.

A combined model with 10 terminal nodes was identified at a relative cost of 0.361. The predictive variables for the model included the average daily tidal range, the maximum number of days without flow over the barrages, the average modelled water level, the maximum daily precipitation, modelled water depth in the previous year, average modelled salinity and the variance in average modelled salinity. This model correctly predicted the original terminal node for 53% of the training data set and 58% of the test data set under cross-validation.

When the various terminal nodes were tested for the distinctiveness of their biological communities (Step 4), ANOSIM analyses were conducted on two subsets of the combined long- and short-term data sets. The first included the presence/absence of bird and commercial fish species ($n = 175$). The second looked at the CPUE of commercially-fished species only ($n = 223$). These analyses, conducted in PRIMER 6.0 suggested that several of the terminal nodes did not support unique biological communities, with no significant difference between two pairs of terminal nodes. These nodes were then combined, and the resulting biota re-assessed using ANOSIM again. The remaining eight terminal nodes did, in fact, represent significantly-different biotic assemblages. As occurred for the short-term model, the two pairs of nodes that were combined were the last two split off from the combined CART model, so creating a model describing the eight (as opposed to 10) terminal nodes was simply a matter of pruning the existing tree, rather than creating a new model. The final model is presented in the results below.

An assessment of the full predictive capacity of the combined ecosystem state model was not straightforward. The misclassification rates reported for the CART analyses apply to the most recent step in the analysis, as they indicate the proportion of cases misclassified with respect to the response variable. For example, in the combined model, this response variable was the state identified either through the long-term or the short-term CART model. This is hardly an independent assessment of the variance in the original environmental and biological data sets, given that it included a mixture of short- and long-term data, and not all sites were included for all years/seasons (depending on the terminal nodes that each were assigned to). In order to provide some information about how well the variance in the environmental and biological data sets could be used to predict the ecosystem state of the Coorong, we undertook a CAP analysis on each of the long- and short-term biological and environmental data sets. CAP analysis allows an assessment of how well values of a response variable (here, ecosystem state) can be predicted using a multivariate data set (here either the biological or the environmental data sets) (Anderson *et al.*, 2008). The analysis has been done from this perspective (rather than using the state to predict the biota or the environmental conditions), as this is the manner in which the model was constructed. CAP analysis was undertaken in PRIMER 6 with the PERMANOVA+ add-on package (Anderson *et al.*, 2008)

The stability of the ecosystem state model was tested using a bagging procedure for the CART tree in CART 6.0 (Steinberg and Golovnya, 2007). This was done using the same variable input predictors as the ecosystem state model presented below, with no cross-validation, but a 0.3 proportion of the data used as a reserve (or holdout) for testing the committee of trees. As occurred for the original model development, cases that had been excluded from the initial clustering due to missing values (at Step 1) were excluded from the bagging routine. The initial tree produced was somewhat different from the model presented here, using the maximum number of days without flow as the initial splitting variable. This initial tree had a prediction success rate of 0.73, with the committee of trees correctly predicting 63% of cases. The committee of trees included 100 trees overall. While the bagging did not produce identical results to the model used here as the ecosystem state model, the committee of trees repeatedly used the same variables included in the ecosystem state model. The most common variation was for the maximum number of days since flow to be the first splitting variable, with the tidal range as the second variable, effectively splitting the data into the same four segments as occurs via the first two splits in the actual

model. Overall, the results of the bagging procedure indicate the structure of the model is quite stable, and that the predictive success of the ecosystem state model is reasonable (53% versus 63% for the committee of trees).

Characterisation of the ecosystem states (Step 6)

Each of the eight identified ecosystem states was characterised based on the biological community they supported and the environmental conditions under which they occurred. These eight ecosystem states were divided into two basins; a marine basin and a hypersaline basin, based on the tidal range which they included. Non-parametric multi-dimensional scaling plots were created from Bray-Curtis similarity matrices of standardised, log(x+1)-transformed abundance data for the biological data sets, and Euclidean-distance similarity matrices based on normalised data for the environmental data sets. A maximum of 25 runs was used for each. ANOSIM analyses were used to test for differences among ecosystem states, and SIMPER to identify species typifying ecosystem states, or distinguishing between them. These analyses were conducted in PRIMER 6.

Mapping the ecosystem states

The ecosystem states of the Coorong were mapped for each of the years over which the model was constructed (1999 to 2007). This involved colour-coding maps of the Coorong according to the identified ecosystem states. Areas between focal sites were coloured to a point half-way between the current site and the next site, with no attempt to identify any zone of transition. This was due to the discrete nature of the ecosystem states (a site can only occur in one at a time), and the lack of data between focal sites to allow a better model to be developed to weight the distance over which each state was likely to occur.

2.1.4.   Developing a predictive model

The intention behind developing an ecosystem state model for the Coorong was that it be used as a tool to predict the likely ecological consequences of a variety of management strategies under a range of climate change conditions. As such, it was necessary to turn the ecosystem model into a predictive S&T model suitable for assessing a range of scenarios.

The various management levers and forcing factors (such as climate) were manipulated to create a number of scenarios (see below for a detailed list). These formed the inputs into the predictive model and drove the physical conditions to determine the ecosystem states at each site for each time step. Analyses of the outcomes of the model runs were then used to assess the likelihood of the water benefit objectives being met for that scenario. This process is illustrated in Figure 2.3.
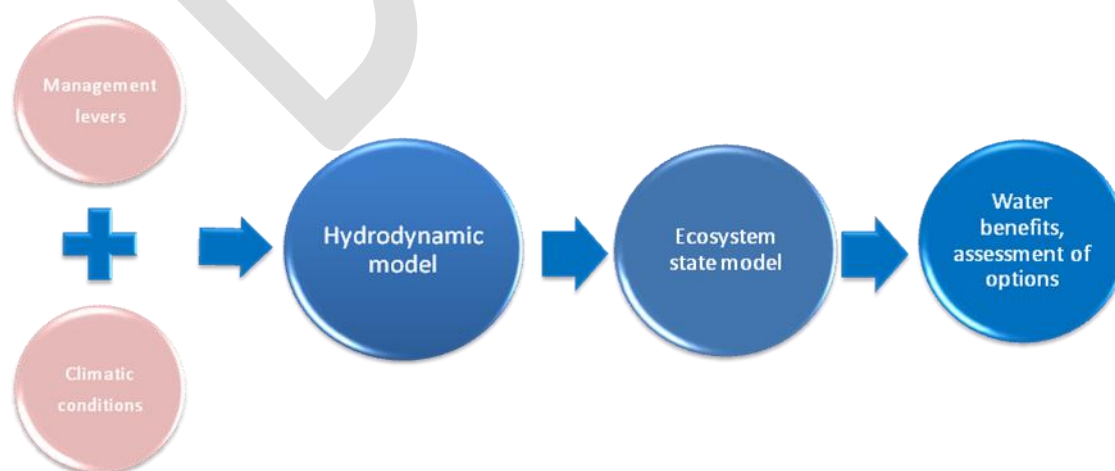
Figure 2.3. Scenario modelling using the synthesised model

By happenstance, all of the drivers of ecosystem condition were ones that could be derived from the output of the hydrodynamic model for the Coorong (Webster 2006). This meant that the largest challenge in predicting the ecosystem states of the Coorong existed for sites outside the domain of the hydrodynamic model. Because of the complexity of the system, and the focus sites of CLLAMMecology as a whole, the hydrodynamic model began at the Murray Mouth and extended in a south-westerly direction (Webster 2006). As such, two of the twelve CLLAMMecology focal were outside the model domain (Monument Road and Mundoo Channel; see Figure 1) and a third was on the edge of the model within the cell simulating the Murray Mouth (Barkers Knoll; Figure 1.1).

For these three sites, it was necessary to extrapolate from the modelled water levels and salinities of nearby cells to predict the environmental conditions occurring at each site. In order to do this, we explored correlations between measured data for each driving variable (e.g. tidal range) for the site of interest with other sites that were wholly within the model domain. The Solver add-in to Microsoft Excel 2007 was used to maximise linear relationships between the variables, exploring uni- and multivariate relationships for the closest site within the model domain (Ewe Island) and several others nearby. This was done by minimising the sum of squared residuals for each relationships through changes to the slope and intercept values. The relationship with the minimum sum of squares over the available data was selected for use in extrapolating modelled data during the scenario analyses. A separate relationship was developed for each variable at each location.

Once the relationships were established, the ecosystem state model was coded as a decision tree in Python 2.6.1 (Python Software Foundation, 2008). The code simply assigned each site in each year in a given scenario to an ecosystem state, based on the values of the variables identified by the model. A time series of ecosystem states along the Coorong was produced as output from the model, which was then post-processed to assess the relative effects of each scenario.

2.1.5. Creating a model independent of barrage flows

Several of the scenarios of interest (see below for full list), included investigating interventions that were intended to have ecological benefits for the Coorong via mechanisms other than changing the flow regime over the barrages. The structure of the ecosystem model (with the maximum number of days without flow over the barrages as the second splitting variable on both sides of the model) meant that any assessment of changes in the distribution of ecosystem states arising from those interventions was unlikely to adequately capture the full extent of change.

To create an alternative model that was independent of flows over the barrages, we used the same eight ecosystem states identified in the original model. We then re-applied the model development procedure without any of the environmental parameters relating directly to barrage flows. This meant that the environmental data set now included data on modelled salinities, water levels, depths and meteorological conditions within the Coorong. The new model was constructed using the combined data set for both long- and short-term environmental data that was used to combine the long- and short-term models above.

The best results were obtained when the two basins identified in the original model were modelled separately. This resulted in two alternative models, one for the marine basin and one for the hypersaline basin. This would diminish the accuracy of the model over the longer term, as it requires an *a priori* assessment of the likely tidal influence at each location, but the interventions designed to mimic the ecological benefit of barrage flows tend to be short-term responses to drought, in particular, so this was not seen as a major problem.

As for the original ecosystem state model, all of the parameters identified as driving the ecosystem states of the Coorong in the alternative model could be calculated from output from the hydrodynamic model. Hourly water levels and salinities along the length of the Coorong that are the output of the hydrodynamic model were used to calculate the average water levels, depths and salinities as required by the ecosystem response models. This allowed us to predict the mixture of ecosystem states present in the Coorong each year for the duration of the model run at each of the 12 focal sites, as occurred for the original model.

In order to test the concordance of the predictions based on the alternative model with those based on the original model, the Baseline scenario (Historic climate with current extraction levels, average inflows from USED and no dredging, see the Scenario Analyses section for additional detail) was run for each over 114 years for each of the 12 focal sites. The proportion of site-years predicted to be in each state was calculated, as was the proportion of degraded states (i.e. those where the number of days without flow has exceeded 339 days) versus the proportion of healthy states (i.e. those with more recent flow over the barrages). The number of states allocated to the marine versus hypersaline basins for the original model was compared with those of the alternative model (where the division between basins was selected *a priori* based on location). The overall concordance of predictions for all site-years was also calculated.

2.1.6. Creating spatial and temporal links in ecosystem states

The initial intention for the ecosystem state model was to incorporate spatial and temporal links between the ecosystem states. We intended to do this by constraining the transitions that were possible in the system through use of an allowable transition matrix. The rationale for this was that the sites would be more likely to occur in a state if that state occurred nearby, either in space or time. This is due to the proximity of propagules, colonists and other refugia making it more likely that the biota associated with a new state would be able to migrate to the location in question. A second reason for incorporating such a matrix was that we considered it likely that some transitions may take longer to occur than others. For example, recovery of ecosystems after degradation would be dependent on the re-colonisation of areas from nearby refugia (e.g. for macrophytes like *Ruppia tuberosa*). This may mean that transitions from more-degraded to less-degraded states may require environmental conditions to be favourable for some time before the change in biota is obvious.

We attempted to constrain the model to only include transitions that had occurred previously (based on a Markov Chain), in the relative proportions that were observed over both the long- and short-term data sets. This was problematic, because the period over which the model was constructed was a period of decline in ecosystem condition, with no recovery occurring. This meant that the probability of recovery ever within the system, according to our modelling, would be zero; a situation that was likely to be incorrect, unnecessarily alarming and not very useful.

As a result, we abandoned the idea of including a transition matrix in the modelling at this stage. It remains an aspect of the modelling approach that we would like to revisit, when additional data are available, particularly regarding emerging recovery pathways.

2.1.7. Model evaluation

We attempted to evaluate the model predictions by comparing the biota from the ecosystem states predicted in the past with the available data for those locations and years. The rationale for this comparison was, if the ecosystem state model can adequately predict all the ecosystem states for the Coorong, we should be able to predict the biotic assemblage of

particular locations and times with accuracy. This is the equivalent of comparing the predictions from a deterministic model to measured data for times outside the calibration period (as was done for the hydrodynamic model; Webster, 2006). The specific hypothesis being tested was that there was no difference between the biota found in any ecosystem state for the years in the training data set (1999-2007) when compared with the biota in the same ecosystem state for data collected prior to 1999.

There was relatively little biological data available with which to evaluate the predictions of the ecosystem state model. Fisheries catch per unit effort data was available from 1984 to 2007. Five years' worth of invertebrate sampling had been undertaken in the early 1980s and repeated in the early 2000s on a presence/absence basis (Geddes and Butler, 1984; Geddes, 1987). This sampling also noted the presences of aquatic macrophytes (including *Ruppia megacarpa* and *Zostera* sp., for example), where relevant. *Ruppia tuberosa* sampling (1984, 1990-1993), along with sampling for bird abundance (1984 and 1985) was undertaken in the South Lagoon of the Coorong (Paton *et al.,* in press). Independent bird sampling was undertaken along the length of the Coorong for 1981, 1982 and 1987 (Gosbell and Christie, 2004).

Based on these data, we compiled data files describing the biological condition of the Coorong for each taxonomic group (birds, fish, aquatic macrophytes and invertebrates) separately across the years covered by the available data. We used the Baseline scenario (Historic climate with current extraction levels, average inflows from USED and no dredging, see Scenario Analyses section below for additional detail) to approximate the environmental conditions at the relevant points in time, and to predict the ecosystem state of each site/region (as appropriate). Where the data were collected at regions spanning more than one CLLAMM focal site (see Figure 1.1), the most commonly-occurring ecosystem state was used as the predicted ecosystem state for that region for that point in time. nMDS ordination plots were constructed based on Bray-Curtis similarity matrices for each taxonomic group. Data had been log-transformed (and standardised if a mixture of sampling methods were used), except for the invertebrate data which were collected as presence/absence only. Ecosystem states identified in the 1980s or in the 1990s were compared with those that were collected during the training data set for the model (1999-2007) using specified contrasts in PERMANOVA (Anderson *et al.*, 2008). The model was structured with ecosystem state as a fixed factor and year as a random factor nested in state. The contrasts investigated were all years from the 1980s compared with years from the training data set (1999-2007), and all years from the 1990s (excluding 1999) compared to the years from the training data set (the latter comparison was only included where data from the 1990s existed). SIMPER analyses identified the taxa contributing most to differences between states or years, and PERMANOVA tested the differences between factors. All analyses were conducted in PRIMER 6 with the PERMANOVA+ add-on software (Anderson *et al.*, 2008).

Separate treatment of each taxonomic did not result in good agreement between data from the 1980s and 1990s and the model predictions, but neither did it provide good agreement between model predictions and the training data set (see below in the Results section). That is, we expected to see clear groupings for each ecosystem state, regardless of the year in which the data were collected, which was largely not observed. This was in contrast to results obtained when combining taxonomic groups, as occurred during the model development. As a result, we also investigated the data available from the 1980s for multiple taxonomic groups where there was overlap between the data sets. This resulted in a very limited number of samples, but we were able to construct nMDS plots for a combination of commercial fish CPUE, aquatic macrophytes and invertebrates, one for a combination of fish CPUE and aquatic macrophytes and one combining fish CPUE and bird data (all based on Bray-Curtis similarity matrices constructed from presence/absence data). PERMANOVA was used to test differences between ecosystem states overall in each comparison, as well as testing the difference between states occurring in the 1980s and

those occurring in the training data set. No multi-taxonomic-group comparisons were possible for the 1990s because commercial fisheries and *Ruppia* data were the only data available.

There were also little data that had been collected more recently than the construction of the model (i.e. in 2008). Fisheries-independent fish abundance data had been collected during 2008 (Noell *et al.*, 2009), but no other data were available. As only a single year was collected, this data was treated using the quarterly time-step, and was assessed against the training data set and model predictions as occurred for the 1980s.

2.1.8. Evidence for additional ecosystem states within the Coorong

While the ecosystem state performs well in describing the ecosystem states that have occurred in the nine years for which we had sufficient data, we acknowledge that other states are likely to (at least potentially) exist that are not adequately represented within this time frame. One that we have identified as likely to occur is an estuarine state, potentially requiring significant, ongoing freshwater inputs, such as have not occurred during the previous decade. Another is a state even less speciose than the degraded hypersaline state in the hypersaline basin, or than the degraded marine state in the marine basin. The existence of both of these states is hinted at in anecdotal accounts of the system, either from the general community or researchers who have worked in the system for many years, and from the trends in data collected after the development of these models, particularly in the South Lagoon. The possible existence of other states that fall outside the bounds of the data set is important to keep in mind when interpreting these results with a view to management of the system.

During the model verification process, a number of PERMANOVA results indicated that there were significant differences between ecosystem states in the 1980s and their counterparts in the training data set.   For this time, most of the ecosystem states were only represented in one or the other time period, and often by a single case. However, the Estuarine/Marine state was one that was predicted for several regions in both the 1980s and in the training data (from 1999-2007). The Estuarine/Marine state was one that we suspected would prove to be a composite of multiple ecosystem states if additional data were available.

We used the three biological data sets that were developed for model verification using multiple taxonomic groups to assess the likelihood that the Estuarine/Marine state as it appears in the ecosystem state model was, in fact, more than one distinct state (presence/absences of commercial fish species and birds, presence/absence of aquatic macrophytes and commercial fish species and presence/absence of aquatic macrophytes, commercial fish species and invertebrates). We cluster analysed each using group-average clustering with a SIMPROF test to identify significant clusters using PRIMER 6. Three significant clusters were identified in both the macrophytes and fish, and macrophytes, fish and invertebrate data sets. Two significant clusters were identified in the fish and birds data set, however, a third cluster could also be identified, and membership within the three was significant when assessed using ANOSIM. The three-cluster solutions for the first two data sets were also significant when tested using ANOSIM.

An environmental data set was compiled from modelled water level and salinity data, modelled flows over the barrages and measured meteorological data. This included all of the same variables that were constructed from the same sources for the original environmental data set during the ecosystem state model development process. The three different cluster solutions were also combined in a similar manner to that used for the short-term data set in the original model development process. There was no overlap between two of the clusters, so four groups were finally allocated (under the assumption that CART and ANOSIM analyses would detect non-significant groups in the response variable). Each

of the cases was assigned to a terminal node based on the CART model, and then the biological community for each was tested using ANOSIM.

2.2. Sensitivity analyses

Several steps within the model development were identified as critical. These were steps where a high degree of sensitivity could dramatically affect the model produced, as well as the interpretation of the condition of the Coorong. The steps identified as critical to the model development process were the:

- initial clustering;

- choice of thresholds; and

- the probability of transitions from one state to another (Lester and Fairweather, 2009b).

The possibility of additional states that may exist for the Coorong, but not be represented in the training data set, was also highlighted as a possible area in which the model may be sensitive to the initial data set (Lester and Fairweather, 2009b).

2.2.1. Assessing sensitivity to initial clustering

The initial clustering step was identified as the most critical step within the model development process (Lester and Fairweather, 2009b). High levels of sensitivity to this step had the potential to dramatically alter the model as a whole, because this was the first point at which the preliminary states were identified. There was the possibility that the model may be sensitive to the method of clustering that was selected, that the clustering may be sensitive to small changes in the measured abundance of various biota, or that these changes may propagate through the model to affect the selection of predictive variables and the threshold values within the CART analysis.

The clustering in the model development process used group-average linkage with no predetermined number of clusters to be assigned. An alternative choice at this stage would have been to use k-means clustering, thus to specify the number of clusters produced. We assessed the degree of sensitivity of the long-term and short-term data sets to the use of group-average linkage clustering versus k-means clustering by undertaking both analyses on each data set and identifying the level of concordance in the case membership of clusters produced by each method. For the k-means analysis of each data set, we set k equal to the number of clusters identified in the original model development (that is, $k = 3$ for the long term data set, $k = 4$ for the short term data including juvenile macroinvertebrate abundances and $k = 5$ for the short term data including fisheries-independent fish CPUE). The k-means analyses were undertaken in SYSTAT 11, while the group average clustering was undertaken in PRIMER 6. The level of concordance for random numbers generated with the same number of possible groups was also produced for comparison.

In order to assess the effect of small changes in the abundances of biota on the configuration of the clusters identified, we systematically changed increasing proportions of the original biological data sets. In sequence, we randomly replaced 1%, 2%, 5%, 7%, 10%, 15%, 20%, 30%, 40% and 50% of the original data and ran a k-means cluster analysis on each new data set, assessing the concordance between the case-membership of the original clusters and that of the altered data sets. This was achieved by replacing a randomly-selected value with a new value, sampled at random from a normal distribution with the same mean and standard deviation as the variable that was being altered. This was

repeated until the desired proportion of the original data had been replaced. Data were then log(x+1)-transformed and a Bray-Curtis dissimilarity matrix was produced. *K*-means clustering was chosen to avoid problems comparing analyses with different numbers of clusters. Each data set was investigated separately (long term, short term with juvenile macroinvertebrate abundances, and short term with fisheries-independent fish CPUE), and the analysis was repeated 1000 times. These analyses were undertaken using R 2.8.1 (www.r-project.org).

A random selection of five runs with 20% substitution and 50% substitution within the long-term data set were reserved in order to assess the sensitivity of the CART analysis to small changes in the original clusters. The *k*-means solution for the long-term data set was also used for this purpose, and compared with the original long-term model using the group-average clusters. For each cluster solution, a CART analysis was run, using the same methods as described in the original model development. The resultant terminal nodes were used as factors in ANOSIM analyses in PRIMER 6 to test for biological differences and where there was no significant difference, the nodes were combined. A final CART analysis was run for the resultant states from the *k*-means clustering and one of the five clusters where 20% of the biological data had been randomly changed. These models were then used to run the Baseline scenario (Historic climate with current extraction levels, average inflows from USED and no dredging, see the Scenario Analyses section for additional detail) to predict the states for each site-year in that scenario. The proportion of site-years with concordant states was then calculated.

A second alternative would have been to base the initial clusters on the environmental data sets, instead of the biological data set. For the long-term data set we analysed the environmental data using group-average clustering in PRIMER 6. This produced a large number of significant clusters according to the SIMPROF test in PRIMER 6, so we also clustered the long-term data using *k*-means clustering in SYSTAT 11. For the short-term data set, we compared *k*-means clustering ($k = 7$) of the environmental data to the amalgamated clusters in the biological data to assess the concordance in the cases assigned to each cluster.

2.2.2. Assessing sensitivity to threshold values

For the predictive model, sensitivity to the exact values identified for each threshold had the potential to affect the results and interpretation of scenario analyses. We assessed the likelihood that the exact value chosen for each of the thresholds was a result of chance alone. In order to do this, we replaced the value of each threshold with a random number drawn from a normal distribution with the same mean and standard deviation as the variable in question from the Baseline scenario. We compared the states predicted for each site-year under the Baseline scenario to a uniform distribution of states. This was done for each threshold in turn and for all thresholds simultaneously, with each repeated 1000 times. Histograms of the resulting distributions were plotted, and the number of runs where the result was as extreme as, or more extreme than, the observed outcome was calculated. Dividing this by the number of runs gave the probability that the observed outcome could be attributed to chance alone.

We also tested the resultant state distribution relative to the distribution observed under the ecosystem state model. We again replaced each threshold with a random number drawn from the distribution of that variable. This time, we calculated the change in the number of site-years predicted to be in each state relative to the original model as a proportion. This analysis was again repeated 1000 times for each threshold individually, and also for all thresholds changing together. The mean proportion of site-years predicted to be in a different state and the coefficient of variation were calculated to give an understanding of how sensitive the resultant states were to the exact values for the thresholds.

2.3.   Scenario analyses

In order to assess the likely ecological outcomes of climate change and to potential management actions, the predictive model was applied to a set of 20 possible future scenarios for the Coorong. The selected scenarios were based on a series of workshops, meetings and discussions with natural resource managers and other stakeholders. These meetings included the second CLLAMM Futures workshop in which possible options were prioritised to produce a short list (Lester and Fairweather, 2008a). This short list was then circulated for comment and finally adopted for CLLAMM Futures. Resource and time constraints dictated that 20 was the maximum number of scenarios that could be attempted during CLLAMM Futures. The 20 that were selected were designed to cover as many key decisions and areas of uncertainty as possible, including the effect of climate change, extraction levels and sea level rise, the effect of The Living Murray initiative, the ecological benefits associated with the Murray Mouth dredging program in times of low flow and the maximum potential benefit of an enhance Upper South East Drainage (USED) scheme. A list of these scenarios is presented below in Table 2.3 and the results of these scenario analyses are presented in (Lester *et al.* 2009).

No.	Scenario	Climate	Extraction levels	Flow over barrages	USED inflows	Mouth dredging	Sea level rise	TLM infrastructure
Benchmark conditions								
1	Baseline	historic (MDB SY Scenario A)	+	+	+	-	-	-
2	Historic Natural	historic	-	+	+	-	-	-
Effects of climate change to 2030								
3	Median Future	median (MDB SY Scenario B)	+	possible	+	-	-	-
4	Dry Future	dry (MDB SY Scenario C)	+	possible	+	-	-	-
5	Median Natural	median	-	possible	+	-	-	-
6	Dry Natural	dry	-	possible	+	-	-	-
Effects of sea level rise								
7	Median Future, -10 cm SLR	median	+	possible	+	-	minimum (10 cm decrease)	-
8	Medium Future, +20 cm SLR	median	+	possible	+	-	median (20 cm rise)	-
9	Median Future, +40 cm SLR	median	+	possible	+	-	high (40 cm rise)	-
10	Dry Future, -10 cm SLR	dry	+	possible	+	-	minimum (10 cm decrease)	-
11	Dry Future, +20 cm SLR	dry	+	possible	+	-	median (20 cm rise)	-
12	Dry Future, +40 cm SLR	dry	+	possible	+	-	high (40 cm rise)	-

Table 2.3. Summary of scenarios investigated as a part of CLLAMM Futures and presented in this report

Note: '+' denotes current levels or present in the scenario and '-' indicates none or not present in the scenario

No.	Scenario	Climate	Extraction levels	Flow over barrages	USED inflows	Mouth dredging	Sea level rise	TLM infrastructure
Effects of TLM initiative								
13	Historic TLM off	historic	+	+	+	-	-	present but no added 500 GL
14	Historic TLM on	historic	+	+	+	-	-	present, 500 GL flows
15	Median TLM off	median	+	possible	+	-	-	present but no added 500 GL
16	Median TLM on	median	+	possible	+	-	-	present, 500 GL flows
17	Dry TLM off	dry	+	possible	+	-	-	present but no added 500 GL
18	Dry TLM on	dry	+	possible	+	-	-	present, 500 GL flows
Effects of other management interventions								
19	MM Dredging	historic	+	+	+	+ (-2m depth)	-	-
20	Max USED Flows	historic	+	+	maximum possible	+	-	-

Table 2.3 cont. Summary of scenarios investigated as a part of CLLAMM Futures and presented in this report

Note: '+' denotes current levels or present in the scenario and '-' indicates none or not present in the scenario

The data available for use during the modelling exercise was a consideration during the development of the list of scenarios. In particular, we wished to align this modelling with that undertaken elsewhere in the Murray-Darling Basin, and so were keen to make use of the CSIRO Murray-Darling Basin Sustainable Yields (MDB SY) project (CSIRO, 2008). The MDB SY project included climate and river modelling for three scenarios that were of interest within CLLAMM Futures. Scenario A was the historic climate sequence with current levels of extraction applied for the duration of the model run. Scenario B included a median projection of 2030 climate, again with current extraction levels operating over the entire model run. Scenario C was the equivalent with a dry climate projection. These scenarios had been modified slightly by the MDBA in order to take into account more-recent climate projections for Victorian rivers for use during the modelling of The Living Murray scenarios (J. Davis, MDBA, pers. comm.). Flows at the barrages based on the modified version of each scenario were used as the input data for modelling the ecosystem states of the Coorong. This information was converted into the maximum number of days without flow by creating a counter that reset each day that flow occurred over the barrages. The maximum for each time step was the highest number that was reached during that period, without resetting the value to 0 at the beginning of each (and so, could therefore be higher than 365 for the annual time step, for example).

Parameters other than flow were derived from the hydrodynamic model of the Coorong (Webster, 2006). The hydrodynamic model produced a time series of water levels and salinities along the length of the Coorong at a daily time step. For each day, maximum, minimum and average water levels, as well as average salinities were produced for each cell. These were then converted to daily averages for the water level (m AHD) and water depth (m) and salinity (g L^{-1}). Tidal range (m) was calculated for each day by subtracting the minimum water level from the maximum. This means that not all of the variation in water level is due to tidal movement, wind seiching and other factors may also play a part, but the variable is referred to as 'tidal range' for simplicity. An annual (or a seasonal) average of the daily values was then calculated for each variable and used as input for the scenario modelling.

The ecosystem response model produced a time series of ecosystem states for each of the 12 CLLAMMecology focal sites (Figure 1.1). This time series was analysed in a number of ways to elucidate the ecological response of the Coorong to each of the scenarios investigated. The 114-year time series was chosen as it used all the available river-flow data that was simulated by MDBA.

For each scenario, a number of analyses were undertaken to elucidate the hydrodynamic and ecosystem state characteristics of the output. Boxplots were constructed for each hydrodynamic variable driving ecosystem states (i.e. tidal range, maximum number of days without barrage flow, water level, water depth from the previous year and salinity for the original model). These displays were undertaken using SYSTAT 11.

Threshold analysis was also undertaken for each driving hydrodynamic variable for each scenario. The years in which each site exceeded the threshold for each variable were selected, and the return time and duration of exceedance were calculated. The return time was calculated as the number of years between when a site fell below the threshold and when the site exceeded the threshold once again. The exceedance duration was simply the number of years for which a site exceeded a particular threshold for each event, and was averaged when the threshold was crossed multiple times. These values were calculated for each site and basin.

The proportion of site-years exceeding salinity of 100 g L^{-1} was also calculated for each scenario. A salinity of 100 g L^{-1} is considered to be important by many experts as an upper tolerance limit for key species in the region (D. Paton, M Geddes, University of Adelaide, pers. comm.; Lester *et al.*, 2008).

A Gini coefficient was calculated for each driving hydrodynamic variable for each scenario. Gini coefficients have only recently been applied to ecological problems (Naeem, 2009; Wittebolle *et al.*, 2009) but are more commonly used in economic analyses, as a measure of the evenness of incomes within a population. The Gini coefficient is thus a measure of evenness of dispersion of a variable. It is calculated by finding difference in the area under a curve based on a cumulative distribution (here, of a hydrodynamic variable) and a Lorenz curve. Values vary between 0 and 1, with 0 representing a perfectly evenly-dispersed distribution and 1 representing a completely unevenly-dispersed distribution.

The deviation from the Baseline scenario was calculated for four driving hydrodynamic variables in each of 20 scenarios. Scenarios were divided into site-years falling above or below the tidal range threshold. For those site-years below the tidal range threshold, the sum of the deviance of each site-year from the Baseline scenario was calculated for the water level and days without barrage flow variables. For site-years over the tidal threshold, the sum of deviance was calculated for salinity and days without barrage flows. These variables were based on the ecosystem state model being used. For the marine basin, depth in the previous year was also a driving variable, but inspection of scenario results indicated that it was only exceeded in a limited number of sites, so for ease of presentation and interpretation, the analysis was limited to two variables per basin. Analyses were conducted in the R statistical environment (www.r-project.org).

The prevalence of each ecosystem state in each scenario was calculated as the proportion of site-years predicted to be in each state. A Markov Chain matrix was also assembled for each scenario, investigating the frequency with which site-years changed between each set of ecosystem states (Horn, 1975). This was used to calculate the frequency of different transitions, including the frequency of transitions between basins, to more-degraded states, or to less-degraded states. States were defined as degraded if they were only present when the Coorong had not received any barrage flows in 339 or more days.

Runs analyses tested the sequence of states appearing at each site in each scenario (Zar, 1999). A runs analysis assesses the significance of a given time series of categorical variables, and compares it to a random time series of the same number of categories. Analyses were conducted in the R statistical environment.

Finally, the ecosystem states of the Coorong were mapped for a set of seminal years (including wet and dry years and the year of Ramsar listing; thus 1902, 1956, 1974, 1985, 2005 and 2008). This involved colour-coding a map of the Coorong according to the state predicted for each of the CLLAMMecology focal sites. Areas between the sites were coloured according to the state of the nearest focal site. No attempt was made to establish zones of transition or to weight by distance the interpolation between sites.

All results for the various scenarios are presented in Lester *et al.* (2009) and Langley *et al.* (2009), along with comparisons between scenarios and management implications.

3. Results

3.1. An ecosystem state model for the Coorong

The final ecosystem state model for the Coorong identified eight distinct ecosystem states (Figure 3.1).
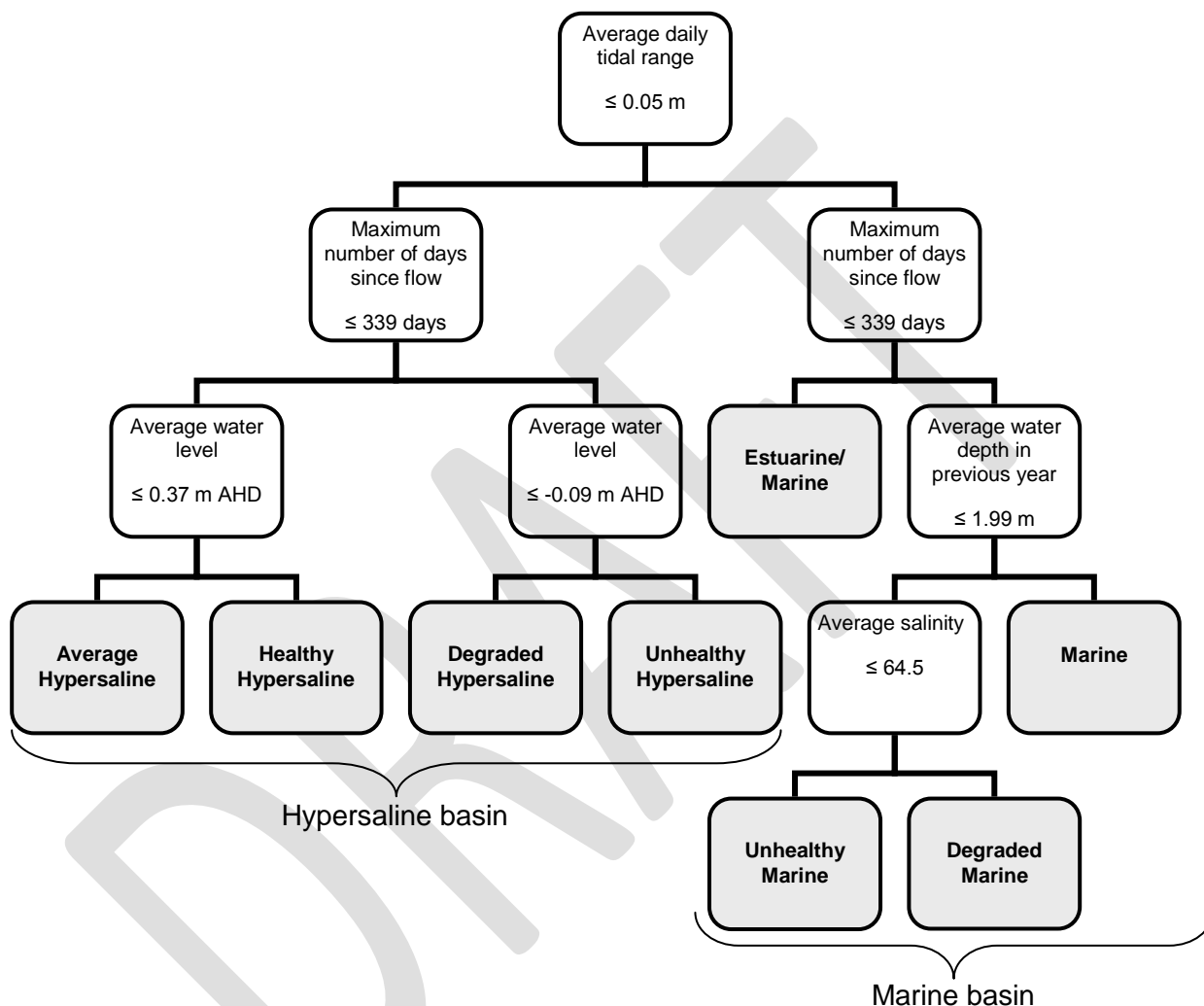
Figure 3.1. Ecosystem states model for the Coorong as a whole

The states are presented as a logic tree, where each box should be read as a logic statement. For a given site-year, if the condition in the box is true, the tree should be followed to the left-hand side. If the condition is false, the tree proceeds to the right, until a shaded terminal node is reached. This terminal node determines which state the Coorong is in at any given location and time, based on its environmental characteristics.

The variables from the environmental data set that predicted the ecosystem state of the Coorong were a combination of water quality, quantity and flow variables. They were average daily tidal range, maximum number of days since flow has crossed the barrages, average water level and salinity, and average depth of water from the previous year. The model was effectively divided into two basins by the appearance of average tidal range as the first split variable. This divided the Coorong into a basin with a moderate to high degree

of tidal influence (called a 'marine' basin) and a basin with little tidal influence (called a 'hypersaline' basin). The marine basin extended from around the Murray Mouth estuary down the northern part of the North Lagoon, to about Noonameena (see Figure 1.1), where the states were influenced by their connection to the Murray Mouth, so they were considered more marine in character (on the right side of Figure 3.1). The hypersaline basin, which included the southernmost part of the North Lagoon and the entire South Lagoon, had four hypersaline states (shown on the left side of Figure 3).

For ease of interpretation, we have given each of the eight states a name (Figure 3.1). The names chosen were based on the environmental conditions under which each state exists, and the range of biota supported by each. The values for each of the thresholds that were significant in the model (and the length of time without water over the barrages, in particular) along with an observed trend of declining biotic richness across a number of the states led us to believe that the ecosystem states represent a continuum from a healthy ecosystem to a degraded ecosystem in each basin. Thus we have named the states accordingly. The names, however, are not intended to identify a single state (e.g. the 'healthy' state) that should occur within a basin. It is not even necessarily possible that such a condition could occur. One of the unique characteristics of the Coorong is that it has always supported a wide range of biota (Phillips and Muller, 2006). The ideal condition for the system as a whole is therefore likely to include a mixture of these states. In addition, the definition of a state used by this modelling approach means that these states may either be transitional states, which would exist for a period of time before changing into another state, or be stable over longer periods. Within the data set available, it is not possible to identify which states are stable over the long term and which are not. Finally, despite the labelling of 'unhealthy' and 'degraded' for several states, this is not to say that no biota exists therein. As is described below, each state continues to support a range of biota that is available as food and habitat resources.

The marine basin consisted of four states, including those named **Estuarine/Marine**, **Marine**, **Unhealthy Marine** and **Degraded Marine**. By definition, these states had greater tidal ranges than those four of the hypersaline basin: **Healthy**, **Average**, **Unhealthy** and **Degraded Hypersaline**. For each, their long-term biological and environmental characteristics have been compiled (Tables 3.1 and 3.2), as have the short-term characteristics of each state (Tables 3.3 and 3.4).

	Marine states				Hypersaline states			
Species	Estuarine/ marine	Marine[†]	Unhealthy Marine	Degraded Marine[‡]	Healthy Hypersaline	Average Hypersaline	Unhealthy Hypersaline	Degraded Hypersaline
Macrophyte (% cores containing *Ruppia*)								
Ruppia tuberosa	0.02	0.00	0.21	NA	NA	51.97	22.69	NA
Fish (Average CPUE (net days))								
Yellow-eyed mullet	6.08	6.62	8.47	9.45	NA	3.90	0.00	NA
Bony bream	0.70	0.03	1.28	0.00	NA	0.53	0.00	NA
Mulloway	1.02	0.69	0.89	0.11	NA	0.13	0.00	NA
Greenback flounder	0.40	0.07	0.32	0.52	NA	0.09	0.00	NA
Bronze whaler shark	0.22	0.41	0.35	0.00	NA	0.00	0.00	NA
Australian salmon	0.84	1.06	1.23	0.00	NA	0.03	0.00	NA
Birds (Average abundance)								
Australian shelduck	383	0	286	13922	NA	1023	905	NA
Red-necked avocet	160	18	61	1260	NA	506	613	NA
Hoary-headed grebe	147	0	265	497	NA	686	904	NA
Curlew sandpiper	283	21	711	238	NA	288	5	NA
Chestnut teal	1047	3	655	5593	NA	1341	373	NA
Banded stilt	529	0	159	15487	NA	2418	20270	NA

Table 3.1. Summary of long-term biological characteristics of Coorong ecosystem states

Note: Biota included are based on the results of SIMPER analyses outlining the species that drove the similarities within states and the differences between states for the long-term analysis. Several SIMPER analyses were undertaken on various subsets of the data due to missing values. The species included here are those that most commonly and significantly differentiate between two ecosystem states.

[†]While the marine state has few of the bird species listed here, it supported more abundant populations of other species. [‡] Caution should be used in interpreting these results, because only one case for the degraded marine state exists. Figures are presented as indicative only. NA = No sites were categorised as healthy hypersaline or degraded hypersaline in the long-term analyses.

Parameter	Marine states				Hypersaline states			
	Estuarine/ marine	Marine	Unhealthy Marine	Degraded Marine[‡]	Healthy Hypersaline	Average Hypersaline	Unhealthy Hypersaline	Degraded Hypersaline
Flow								
Maximum number of days since flow (days)	127	376	376	376	NA	152	376	NA
Water quantity								
Average daily tidal range (m AHD)	0.11	0.38	0.19	0.00	NA	0.03	0.03	NA
Variance of average water level	0.03	0.02	0.02	0.02	NA	0.08	0.03	NA
Water quality								
Maximum salinity (g L^{-1})	57.56	48.11	61.64	88.24	NA	123.28	176.47	NA
Average [total phosphate] (mg L^{-1})	0.14	0.05	0.12	0.35	NA	0.27	0.50	NA
Average [ammonia] (mg L^{-1})	0.04	0.05	0.05	0.04	NA	0.20	0.30	NA
Average [Chlorophyll b] (µg L^{-1})	2.38	0.45	2.33	4.20	NA	10.39	7.11	NA
Average turbidity (NTU)	14.59	1.68	11.14	35.00	NA	17.55	23.53	NA
Average [TKN] (mg L^{-1})	1.85	0.61	1.22	6.57	NA	5.61	6.90	NA
Meteorological*								
Maximum daily minimum temperature (^{o}C)	24.06	28.75	28.73	29.00	NA	24.11	28.75	NA
Average annual rainfall (mm $annum^{-1}$)	467.46	300.95	301.35	295.80	NA	476.10	300.95	NA

Table 3.2. Summary of long-term environmental characteristics of Coorong ecosystem states

Note: Parameters included are based on the results of SIMPER analyses outlining the variables that drove the similarities within states and the differences between states for the long-term analysis. Several SIMPER analyses were undertaken on various subsets of the data due to missing values. The parameters included here are those that most commonly and significantly differentiate between two ecosystem states. [‡] Caution should be used in interpreting these results, as only one case for the degraded marine state exists. Figures are presented as indicative only. * Meteorological parameters were calculated from data collected at Meningie weather station (see Figure 1). NA = No sites were categorised as healthy hypersaline or degraded hypersaline in the long-term analyses.

Species	Marine states				Hypersaline states			
	Estuarine/ marine	Marine	Unhealthy Marine	Degraded Marine[‡]	Healthy Hypersaline	Average Hypersaline	Unhealthy Hypersaline	Degraded Hypersaline
Fish (Average CPUE (net days))								
Small-mouthed hardyhead	NA	132.30	962.57	708.00	ND	ND	406.40	79.60
Yellow-eyed mullet	NA	477.40	211.87	171.00	ND	ND	0.00	0.00
Australian salmon	NA	118.00	48.57	15.00	ND	ND	0.00	0.00
Sandy sprat*	NA	294.40	155.74	49.00	ND	ND	0.00	0.00
Birds (Average abundance)								
Chestnut teal	36	10	37	16	96	92	71	36
Hoary-headed grebe	31	11	1	0	295	199	416	3
Banded stilt	6	0	1	54	406	868	1361	1894
Red-necked stint	62	152	78	86	23	499	119	211
Grey teal	106	113	67	35	292	157	176	0
Whiskered tern	34	58	72	1545	2	75	600	3
Invertebrates (Average abundance)[†]								
Capitella	8084	3619	10775	ND	0	0	ND	ND
Arthritica	2439	1561	4500	ND	0	0	ND	ND
Simplesetia	1264	1327	3424	ND	0	0	ND	ND
Insect larvae	319	336	177	ND	740	34	ND	ND

Table 3.3. Summary of short-term biological characteristics of Coorong ecosystem states

Note: Biota included are based on the results of SIMPER analyses outlining the species that drove the similarities within states and the differences between states for the long-term analysis. Several SIMPER analyses were undertaken on various subsets of the data due to missing values. The species included here are those that most commonly and significantly differentiate between two ecosystem states. [‡] Caution should be used in interpreting these results, as only one case for the degraded marine state exists. Figures are presented as indicative only. [†] All invertebrate abundances reported here are for juveniles. * *Hyperlophus vittatus*. ND = indicates that no data were available for that species for that particular state.

Variables	Marine states				Hypersaline states			
	Estuarine/ marine	Marine	Unhealthy Marine	Degraded Marine[‡]	Healthy Hypersaline	Average Hypersaline	Unhealthy Hypersaline	Degraded Hypersaline
Flow								
Maximum number of days since flow (days)	141	540	536	652	2	166	559	505
Average flow volume from previous year (GL)	47.70	55.56	57.48	0.00	56.89	45.81	85.48	0.00
Water quantity								
Average maximum water level (m AHD)	0.34	0.39	0.22	ND	0.42	0.15	-0.03	-0.25
Maximum water level (m AHD)	0.69	0.78	0.56	ND	0.58	0.43	0.33	-0.07
Minimum water level (m AHD)	-0.12	-0.27	-0.25	ND	0.21	-0.11	-0.35	-0.40
Average minimum water level (m AHD)	0.13	-0.01	0.02	ND	0.38	0.12	-0.08	-0.29
Average water level (m AHD)	0.22	0.15	0.12	0.04	0.40	0.13	0.04	-0.25
Average daily tidal range (m)	0.21	0.39	0.20	ND	0.03	0.02	0.03	0.03
Change in water level over quarter (m)	0.81	1.05	0.81	ND	0.38	0.53	0.68	0.32
Water quality								
Average maximum salinity (g L^{-1})	49.82	44.12	52.11	71.60	113.03	123.73	128.87	178.87
Average minimum salinity (g L^{-1})	32.05	34.72	39.45	55.20	99.63	95.22	97.39	120.53
Average salinity (g L^{-1})	41.46	37.98	45.04	64.62	105.94	110.78	112.56	151.37

Table 3.4. Summary of short-term environmental characteristics of Coorong ecosystem states

Note: Variables included are based on the results of SIMPER analyses outlining the factors that drove the similarities within states and the differences between states for the long-term analysis. Several SIMPER analyses were undertaken on various subsets of the data due to missing values. The parameters included here are those that most commonly and significantly differentiate between two ecosystem states.

[‡] Caution should be used in interpreting these results, as only one case for the degraded marine state exists. Figures are presented as indicative only. ND = indicates that no data were available for that species for that particular state.

For five of the eight states, the long-term differences in multivariate displays of either community composition and environmental parameters are shown in Figures 3.2 and 3.3, respectively. Only five states are shown because of missing values excluding a sixth state, and two states did not appear within the long-term analyses. For these five shown, however, there were significant differences in the community composition between states (ANOSIM Global $R = 0.402$, $p = 0.001$), based on the macrophyte, bird and commercial fish abundances found, and in the environmental parameters found (ANOSIM Global $R = 0.536$ $p = 0.001$).
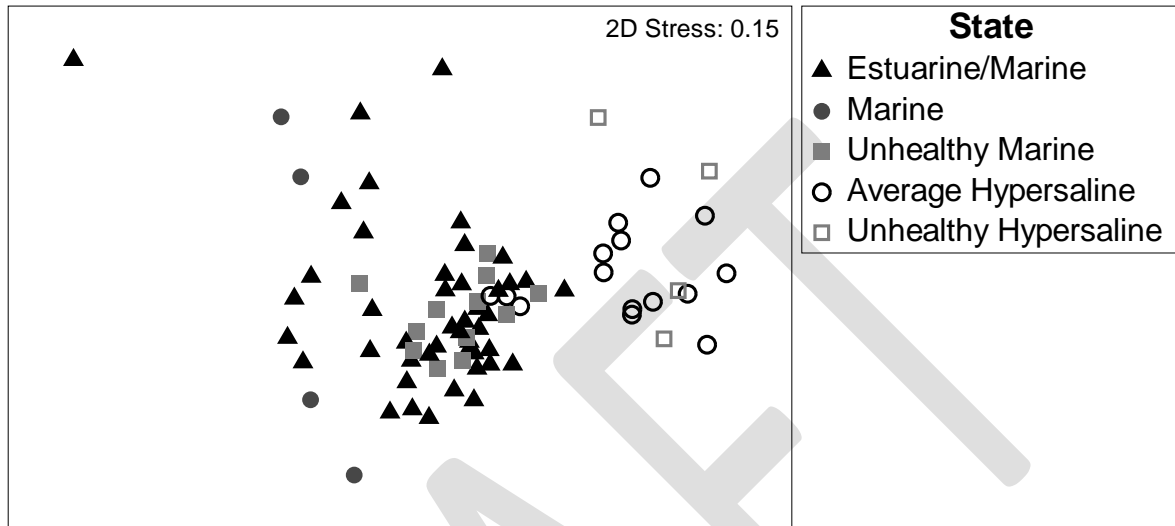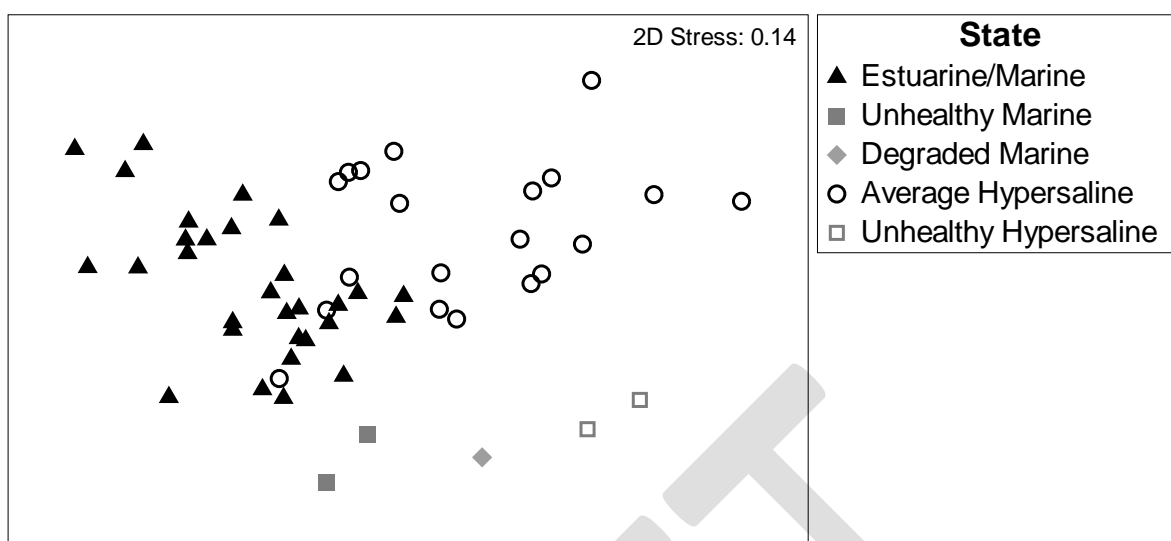

Figure 3.2. nMDS plot of long-term biological characteristics for the ecosystem states of the Coorong.

Note: The following figure is based on a Bray-Curtis similarity matrix of standardised, log(x+1)-transformed abundance data for macrophytes, birds and commercial fish for 12 sites between 1999 and 2007 (minus cases with missing values, $n = 80$). A maximum of 25 runs were used.

This figure shows a non-parametric multi-dimensional scaling plot that give a two-dimensional representation of the similarity between the samples analysed. Each point on the figure represents a single site-year within the Coorong data set. The position of each point is based on its relative similarity to all the other site-years represented. Points that are close together are more similar than points that are further apart. The stress value given in the upper right-hand corner indicates the relative stress associated with representing the multi-dimensional data as a two-dimensional plot. Values of less than 0.2 are widely considered to be acceptable (Clarke, 1993). The points are coded to show the state that each site-year was allocated to, with states within the marine basin shown as solid shapes and states within the hypersaline basin shown as open shapes.

The difference between the biotic composition of each state supports the notion that the model is describing distinct ecosystem states. There is a large distinction between the marine and the hypersaline states, with the closed and open symbols clearly differentiated. The distinctions within these two groups are smaller, with several groups overlapping, indicating that, although statistically significant, there are more similarities within the groups of states. The Estuarine/Marine state, in particular, appears not to be particularly well-grouped. This is not supported by the results of a dispersion analysis using PERMDISP, a test of the dispersion of points within a group (Anderson *et al.*, 2008). There was no significant difference in the dispersion (or relative spread of the points) for the various states. The Marine state was the most dispersed, with the Unhealthy Marine and the Unhealthy Hypersaline state the least dispersed. Pair-wise comparisons showed that the Marine state was significantly more dispersed than the Unhealthy Marine or the Average Hypersaline states ($t = 4.01$, $p = 0.001$ for the Marine state compared to the Unhealthy Marine state and $t = 2.84$, $p = 0.034$ for the Marine state compared to the Average Hypersaline state).

Figure 3.3. nMDS plot of long-term environmental characteristics for the ecosystem states of the Coorong.

Note: The following figure is based on a Euclidean similarity matrix of normalised data for modelled salinity, depth and water levels, flow characteristics, meteorological and water quality data for 12 sites between 1999 and 2007 (minus cases with missing values, *n* = 55). A maximum of 25 runs were used.

The environmental characteristics of all states were much more distinct, with very little overlap between any of the states. This was supported by the pair-wise ANOSIM, which showed that all states were significantly different with the exception of those where only one or two site-years represented that state (Unhealthy Marine and Degraded Marine). Dispersion of points was significantly different for the environmental parameters of each state (*F* = 13.36, *p* = 0.001 overall). Where pair-wise tests were possible (i.e. more than two cases per state), only the Estuarine/Marine and Average Hypersaline states did not have significantly different degrees of dispersion, with these two having the highest overall level of dispersion among the states. The greater level of distinction between the states with respect to their environmental characteristics may be a reflection of the more complete environmental data upon which these analyses is based, and additional biological data may better illustrate the distinction between the various marine and hypersaline states.

The environmental variables performed well in predicting the ecosystem state of the system. CAP analysis on the long-term environmental data resulted in the choice of 11 principal coordinates analysis (PCO) axes (Figure 3.4). Figure 3.4 looks similar to the nMDS plots presented in Figure 3.2 and 3.3, but includes an explicit hypothesis test (that is, that the environmental data set can be used to predict the ecosystem state), rather than simply representing the similarity between individual cases in two-dimensions. In this figure, the position of each point is determined by the value of each case against the first two PCs. The analysis of the long-term environmental data, based on those 11 axes, incorporated 97% of the variance in the original resemblance matrix. Cross-validation based on this analysis correctly predicted the ecosystem state for 93% of cases excluded from the model in turn. Indeed all of the misclassified cases belonged to the Average Hypersaline state being misclassified as Estuarine/Marine. Permutation testing confirmed that the differences among ecosystem states in multivariate space were significant (*tr(Q_m'HQ_m)* = 1.98308, *p* = 0.001).
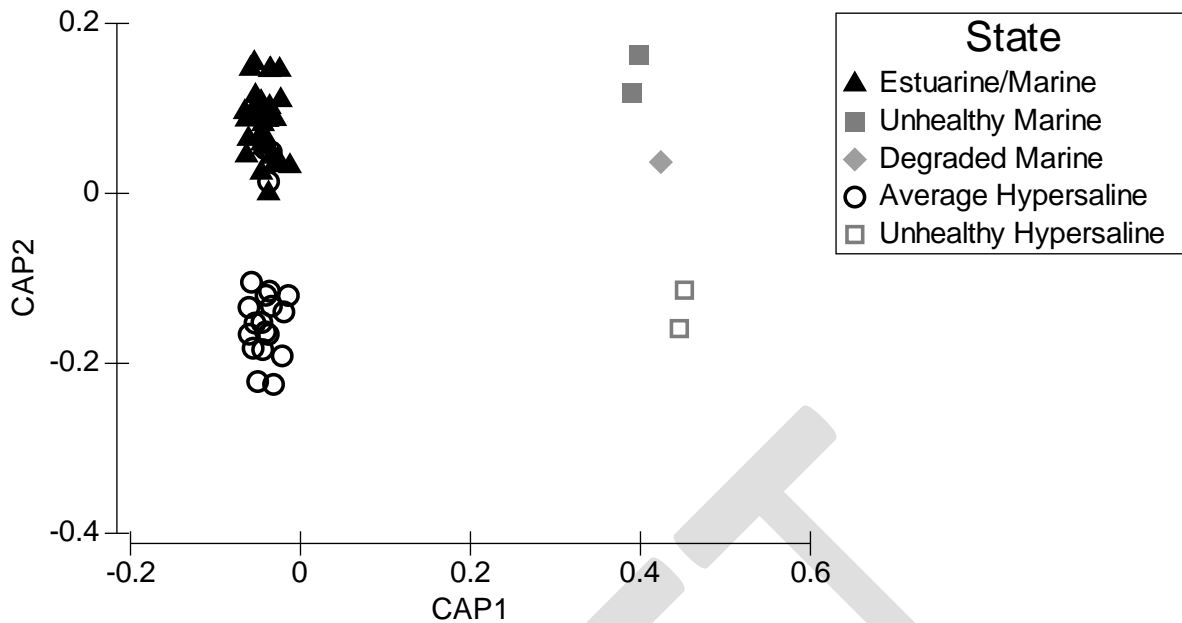
Figure 3.4. Constrained ordination to assess the capacity of the long-term environmental data to predict the ecosystem state of the Coorong

In CAP analyses of the short-term environmental data set, 18 PCO axes were identified, explaining 99% of the original resemblance matrix (Figure 3.5). Here, cross-validation correctly classified 92% of cases excluded sequentially from the analysis. Here, cases were more likely to be misclassified if they belonged to states with few representatives (e.g. Healthy Hypersaline and Degraded Hypersaline). Again, permutation testing confirmed that the differences between groups was statistically significant ($tr(Q\_m'HQ\_m) =$ 4.2356, $p =$ 0.001).
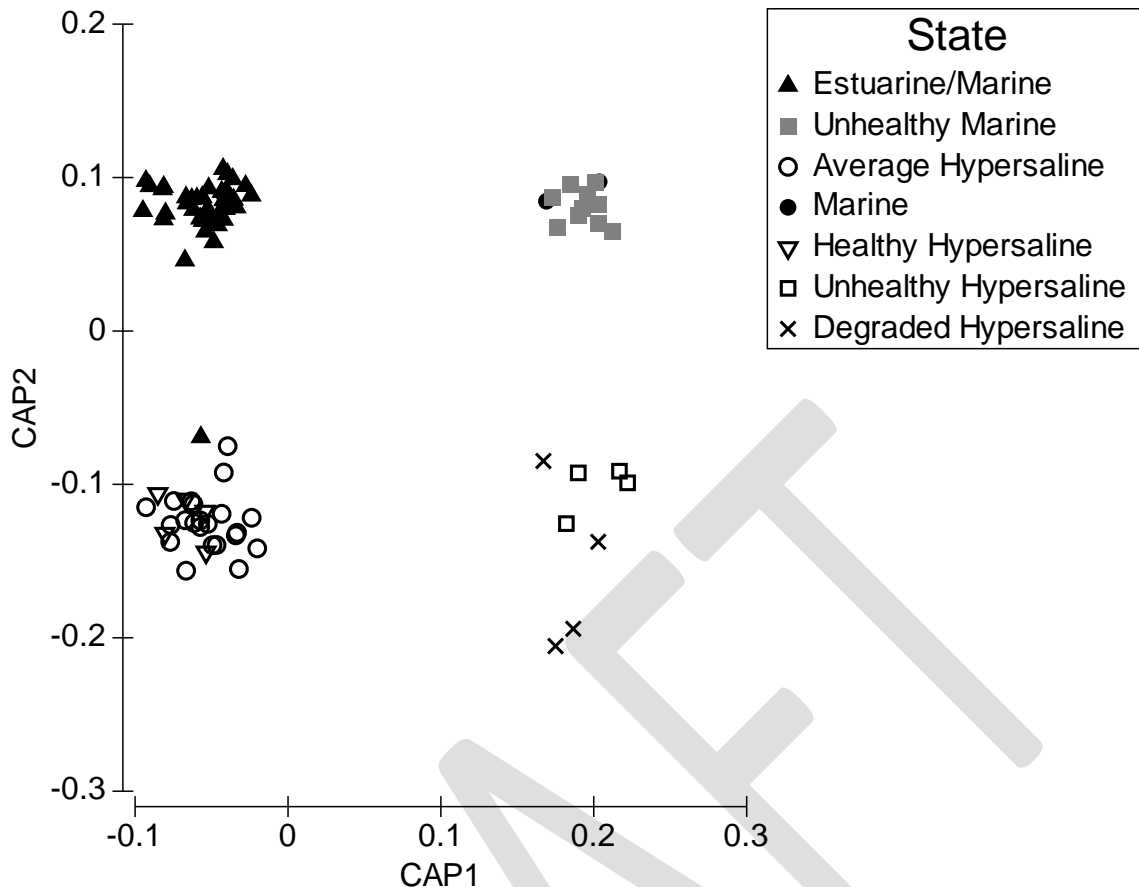
Figure 3.5. Constrained ordination to assess the capacity of the short-term environmental data to predict the ecosystem state of the Coorong

Assessing the ability of the biological data set to predict the ecosystem state (effectively what was done when using the biological data set to identify preliminary states through cluster analysis) had more variable results. For the long-term biological data set, there was a clear ability to predict the relative basin for each case, but less-distinct differences within each basin (Figure 3.6). This was reflected by the misclassification rate (44%) for the selected number of 12 PCOs. The 12 PCOs accounted for 95% of the variation in the original resemblance matrix. Misclassification rates were highest for the Unhealthy Marine and Marine states (66% and 75%, respectively), but there were also relatively few cases (particularly for the Marine state). The analysis had the most difficulty distinguishing between the Unhealthy Marine state and the Estuarine/Marine state, with cases misclassified in both directions. Despite this, permutation testing still identified the differences between states as significant ($tr(Q\_m'HQ\_m) =$ 1.65098, p = 0.001).

Using the short-term biological data set to predict ecosystem states was more successful. Using either the data set including juvenile invertebrate abundances or fisheries-independent fish abundances greatly assisted in resolving the differences between the Estuarine/Marine state and the Unhealthy Marine state.

The CAP analysis including juvenile invertebrate abundances identified 6 significant PCOs, explaining 78% of the variation in the resemblance matrix. This model correctly classified 82% of cases under cross-validation, with Healthy Hypersaline having the highest misclassification rate, but with few cases to classify (n = 5). Permutation testing indicated that differences between ecosystem states were significant ($tr(Q\_m'HQ\_m)$ = 2.128382, p = 0.001).

For the analysis including fisheries-independent fish abundances, 14 PCOs were identified, incorporating 96% of the variance in the resemblance matrix (Figure 3.7). This analysis correctly classified 86% of cases under cross-validation, and permutation testing again confirmed the differences between states as significant ($tr(Q\_m'HQ\_m)$ = 2.55216, p = 0.001).
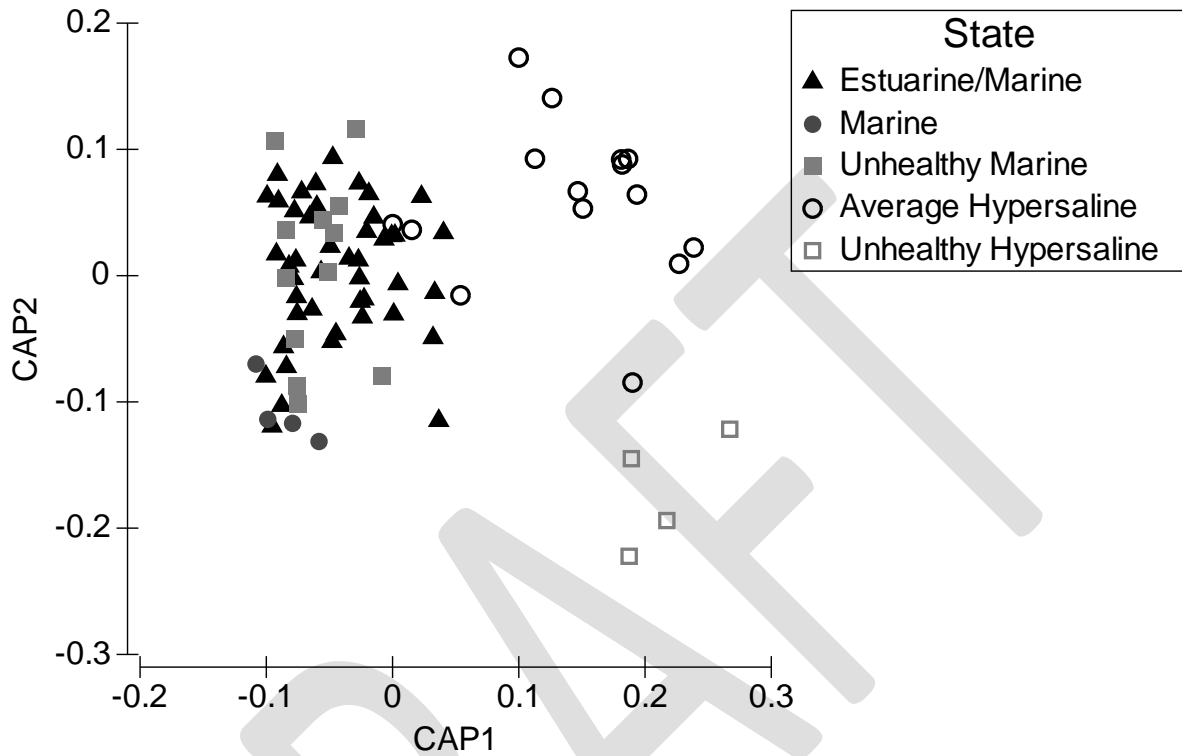
Figure 3.6. Constrained ordination to assess the capacity of the long-term biological data to predict the ecosystem state of the Coorong
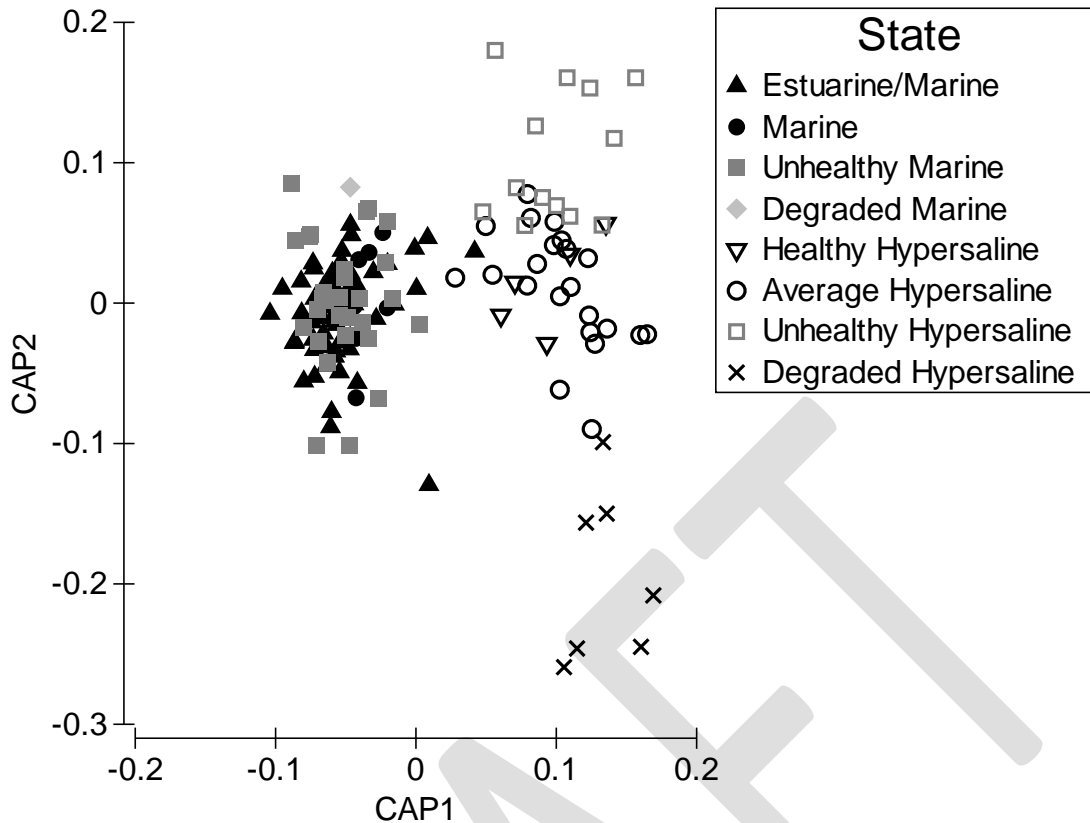
Figure 3.7. Constrained ordination of the to assess the capacity of the short-term biological data to predict the ecosystem state of the Coorong (based on data including fisheries-independent fish abundances)

3.1.1. Characterising each of the ecosystem states

Each of the eight ecosystem states have been described in terms of their biological and environmental characteristics. These narrative descriptions are a combination of the characteristics apparent from both the long- and short-term data sets.

Estuarine/Marine

Affected by tidal influences from the Murray Mouth, the **Estuarine/Marine** state was characterised by lower average salinities and a shorter period since flow occurred over the barrages than other states. This input of both marine and fresh water also led to the state having the highest average water depths and water levels and high variability in water levels across the time period (quarterly for shorter-term analyses and annually for longer-term analyses). In addition, the water quality characteristics of this state may also have reflected both tidal (i.e. via the Murray Mouth) and freshwater (i.e. via the barrages) influences, with low nutrient concentrations (e.g. ammonia and TKN), low chlorophyll (a and b) concentrations and low turbidity. This state supported large numbers of marine and estuarine fish (e.g. yellow-eyed mullet *Aldrichetta forsteri*, mulloway *Argyrosomus japonicus*, greenback flounder *Rhombosolea tapirina*, black bream *Acanthopagrus butcheri* and Australian salmon *Arripis truttaceus*), which were characteristic of this state. The presence of a variety of fish species and tidal influence was reflected in the birds that were present, with the state dominated by piscivores (e.g. several cormorant species *Phalacrocorax* spp. and Australian white ibis *Threskiornis molucca*). There were also other bird groups associated with this state, including waterfowl (e.g. Australian shelduck, *Tadorna*

tadornoides) and migratory waders (e.g. curlew sandpiper) present. The aquatic macrophyte, *Ruppia tuberosa*, a potential food source for some bird species, was also associated with this state, although it was limited in distribution. There were also large numbers of invertebrates that characterised this state, with high abundances of amphipods, *Simplisetia aequisetis* (a polychaete), and adult and juvenile *Capitella* spp. (polychaete taxa).

Marine

The **Marine** state had the lowest average salinity of all the states (although it was not significantly lower than for the Estuarine/Marine state). Like the Estuarine/Marine state, it also had low average water levels, but had the highest variability in water levels across the time period.   Compared with the Estuarine/Marine state, this state was characterised by greater time between water inputs, and more days since flow occurred over the barrages. The water quality may have reflected mostly marine inputs, with the lowest average concentrations of total phosphate, TKN and turbidity across all states. Biologically, this state was dominated by marine and estuarine fish species, including the Australian salmon, bronze whaler shark (*Carcharhinus brachyurus*) and black bream, with fewer yellow-eyed mullet and greenback flounder than the other marine basin states. Fewer piscivorous birds were present (in comparison to other marine basin states), with greater numbers of waterfowl species (e.g. musk duck *Biziura lobata* and pacific black duck *Anas superciliosa*) present. There were also fewer amphipods and capitellids (a family of polychaetes), but greater numbers of *Nephtys australiensis* (another polychaete) and *Arthritica helmsi* (a bivalve).

Unhealthy Marine

As occurred for the Estuarine/Marine and Marine states, the **Unhealthy Marine** state also had relatively low average salinities (but slightly higher than Estuarine/Marine). The average water levels were still high, but there was greater variability in the average water levels across the time period for this state. As for the Marine state, the average maximum number of days since flow occurred over the barrages was higher than for the Estuarine/Marine state, indicating greater time between inputs of freshwater. This lack of freshwater may be reflected by this state having low average depths compared with other states in the marine basin. Average water quality characteristics such as nutrient concentrations (e.g. ammonia, total phosphate and TKN) and turbidity values were low. The Unhealthy Marine state still maintained a diverse fish population, with high abundances of yellow-eye mullet and bony bream (*Nematolosa erebi*). The bird species for this state were dominated by piscivores, including great, little black and little pied cormorant species (*Phalacrocorax carbo*, *Phalacrocorax sulcirostris* and *Phalacrocorax melanoleucos*, respectively) and hoary-headed grebes (*Poliocephalus poliocephalus*). Other prominent species included curlew sandpiper (a migratory wader) and black swan (*Cygnus atratus*, a waterfowl), which may be attracted by the presence of *Ruppia tuberosa* and by the high abundances of adult and juvenile invertebrates, including capitellids, *Simplisetia aequisetis* (both polychaetes) and *Arthritica helmsi* (a bivalve).

Degraded Marine

The **Degraded Marine** state was represented by only a single case in each of the short- and long-term data sets used in the determination of this model (located at Noonameena in the short-term and Parnka Point in the long-term analyses). As a result, we have provided little detail for the characteristics of this state, simply giving some direction as to what this state may represent.   It appears that this state had higher average salinity with lower water levels and inputs. Water quality, although only collected for the long-term data set, appeared to remain low for concentrations of some nutrients (e.g. ammonia) but may be degrading with higher values for others (e.g. TKN concentrations and turbidity). The fauna that was associated with this state included a mix of piscivorous and wading birds, few fish species and chironomid larvae as the only benthic macroinvertebrates present.

Healthy Hypersaline

In the hypersaline basin, the **Healthy Hypersaline** state had higher average salinity values than for the states of the marine basin. This state was characterised by low average variability in water levels across the time period, and with a high average water level. The Healthy Hypersaline state also had high average depths (only lower than the Estuarine/Marine and marine states in the marine basin) and the lowest average maximum number of days since flow occurred over the barrages. Thus, the Healthy Hypersaline state featured frequent freshwater flows, with the average maximum interval between flows being only 11 days; much lower than for all other states. Frequent freshwater inputs were also reflected by this state having the highest average days with flow and the highest average flow volume of all states. With such frequent freshwater inputs, the water quality of this state was characterised by the highest concentrations of nutrients (e.g. ammonia and TKN) and highest turbidity of all states. The Estuarine/Marine state also had high average chlorophyll concentrations. The absence of estuarine and marine fish species (e.g. black bream, greenback flounder and yellow-eyed mullet) may have been due to the higher salinities of this state compared with the northern states, or to high nutrient concentrations, some of which are potentially toxic to fish. The bird fauna was dominated by large numbers of waders and waterfowl, including the grey teal (*Anas gracilis*), black swan, chestnut teal (*Anas castanea*) and red-necked avocet (*Recurvirostra novaehollandiae*). Compared to other states, there were smaller numbers of other wader species and piscivores, including the red-capped plover (*Charadrius ruficapillus*), red-necked stint and whiskered tern (*Chlidonias hybridus*). There were also lower numbers of invertebrates associated with this state, particularly juvenile capitellids, *Simplisetia aequisetis* and *Arthritica helmsi*. However, there were higher numbers of juvenile insects (other than chironomids) compared with all of the other states.

Average Hypersaline

Like the Healthy Hypersaline state, the **Average Hypersaline** state was characterised by higher average salinities than the marine basin. It had moderate change in water levels across the time period analysed (that is, quarterly for the short-term and annually for the long-term analyses). This state had low average depths and received freshwater influences from flow over the barrages reasonably often, with few days between flows. Consistent with freshwater inputs, the water quality indicated high nutrient concentrations (e.g. TKN and ammonia) and the presence of algae and diatoms, with high concentrations of both chlorophyll a and b and high turbidity. The high values of potentially undesirable water qualities (e.g. salinity and ammonia levels) may have been responsible for the lower abundances or absence of various fish species. There were very few fish species associated with this state, with very low numbers of greenback flounder and mulloway and no black bream. Corresponding with the lower numbers of fish, there was also a lack of piscivorous birds associated with this state (with the exception of the Australian pelican, *Pelecanus conspicillatus*). Instead, the bird species associated with this state included other functional groups, including waders (e.g. banded stilt *Cladorhynchus leucocephalus*, red-necked stint and red-necked avocet) and waterfowl (e.g. grey teal). *Ruppia tuberosa* was also more dominant within this state, with greater coverage than any other state (although no data was available for the Healthy Hypersaline state). There were also very few invertebrate taxa associated with this state (e.g. *Capitella* spp., *Nephtys australiensis* and their juvenile equivalents), but chironomid larvae and amphipod species were present in higher numbers than were found at other states.

Unhealthy Hypersaline

The **Unhealthy Hypersaline** state had higher average salinities than the Healthy or Average Hypersaline states. It was also characterised by low average water levels. Despite average depths being mid-range compared with other states, the maximum was relatively low with only 0.4 m difference between the average and maximum depths. Thus, this state was also

characterised with low variability in water levels across the time period. The Unhealthy Hypersaline state had a high average maximum number of days since flow occurred over the barrage, indicating low freshwater flows influencing the state. Like other hypersaline basin states, the water quality indicated high average nutrient concentrations (e.g. ammonia and total phosphate) and high turbidity. Like the Healthy Hypersaline state, such high nutrient levels may have been responsible for the low numbers of fish present, except for the high numbers of small-mouthed hardyhead *Atherinosoma microstoma* present within this state, although high salinities would also significantly affect fish diversity. The low numbers of fish species also led to the relatively low abundances of some piscivorous bird species, including great cormorant and fairy tern (*Sterna nereis*). Some piscivorous species still occurred within the state in reasonable numbers, including the hoary-headed grebe and Australian pelican. Other characteristic bird taxa included banded stilt and Australian shelduck. The state supported a very small diversity of invertebrates (e.g. a few *Simplisetia aequisetis*, *Capitella* spp. and *Arthritica helmsi*), but still had high numbers of chironomid larvae.

Degraded Hypersaline

Like the previous hypersaline basin states, the **Degraded Hypersaline** state had a high average salinity, the highest of all states detected for the Coorong. It also had the lowest average water levels, with a maximum water level of only -0.10 m AHD during the time period. With such low water levels, this state was also characterised by having the lowest average change in water levels over the time period and the lowest average depths. This was likely to be due to the low input of freshwater received from the flow over the barrages, given that this state had the highest minimum number of days since flow (excluding the Degraded Marine state). Attributes of the water quality for this state were variable, with low average ammonia concentrations, but high average total phosphate concentrations and high turbidity (similar to the Average Hypersaline state) and higher average TKN. This state was therefore characterised by more variability in nutrient concentrations, compared with others in the hypersaline basin. The state supported relatively few fish species and was dominated by the presence of small-mouthed hardyhead, which are tolerant of high salinity. Similarly, there was also a lack of invertebrate species, including those known to be more salinity-tolerant (e.g. chironomids). The bird taxa characterising this state were waders and waterfowl, with high numbers of banded stilt and red-necked stint, also with silver gull (*Larus novaehollandiae*), masked lapwing (*Vanellus miles*) and Australian shelduck. Similarly to the other hypersaline states, there were lower numbers of piscivorous species, including the Australian pelican and whiskered tern, and the waterfowl grey teal.

3.1.2. Mapping the ecosystem states in space and time

One example of an application for the ecosystem model that has been developed is that it can be used to investigate changes in ecosystem states in space and over time.

One of the unique features of the Coorong is the range of conditions that it supports at any one time along its 110-km length. This was one of the reasons for its listing on the Ramsar Convention for Wetlands, and its status as a Murray-Darling Basin Icon Site (Murray-Darling Basin Commission 2006). The Murray Mouth region has a tidal influence that is determined by the size of the estuarine mouth opening. The barrages, also toward the north of the system, can provide some freshwater flows, potentially creating estuarine conditions. Further south in the system, the distance from these two sources of water increases, and so the water becomes more saline due to evaporation, and the South Lagoon is often disconnected from (or poorly connected to) the North Lagoon, limiting the source of water to that lagoon. This range of physical conditions has always supported a wide range of biological communities, making a mix of ecosystem states the normal condition for the Coorong. It is important that our modelling reflect that mix. We mapped the changes in

ecosystem state for the Coorong (Figure 3.8), and describe the observed changes in ecosystem response that have occurred over the last 10 years by focussing on the beginning (2000) and the end (2007) of the time series.
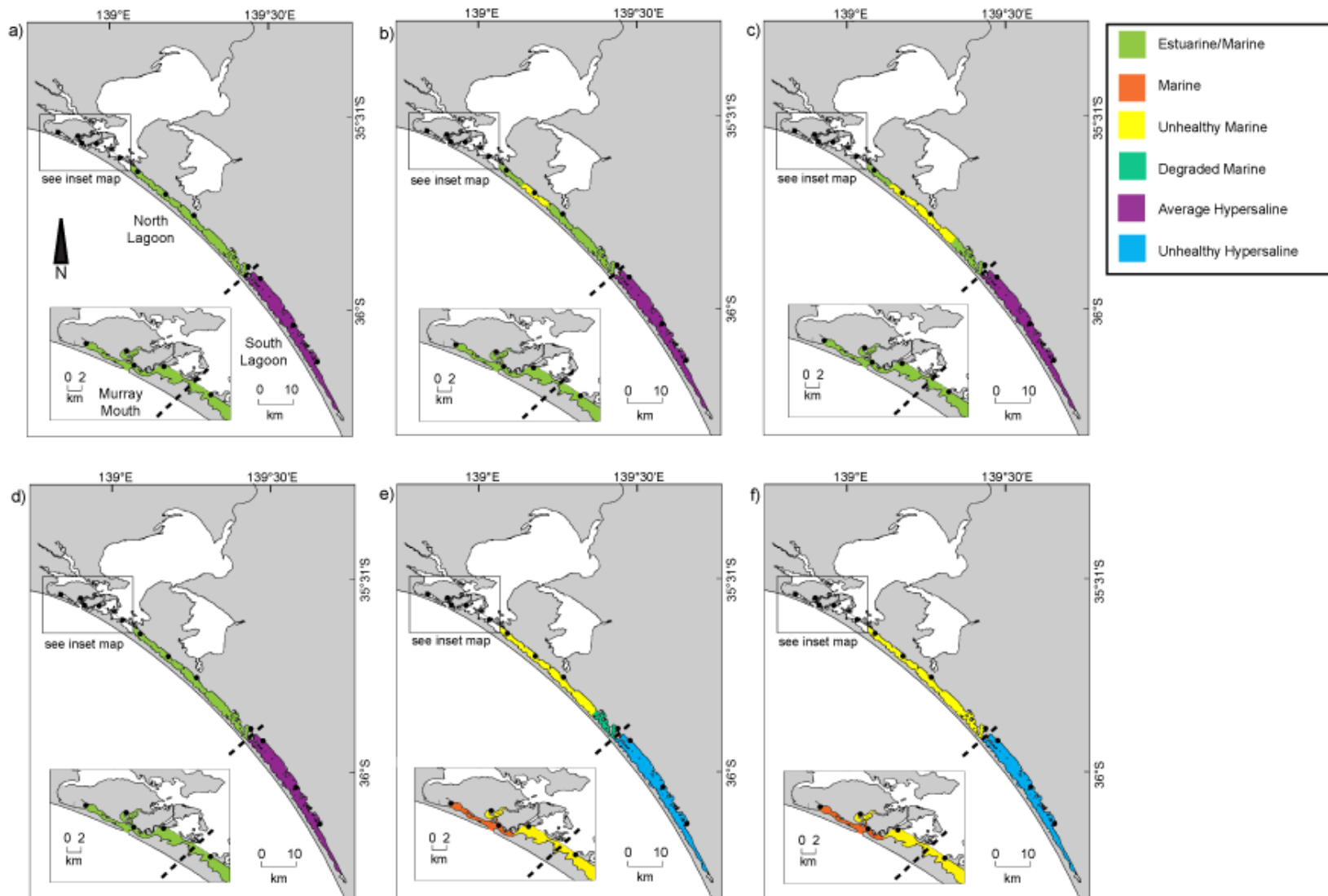
State of the Coorong in 2000

Near the beginning of our time series, during 2000, two ecosystem states were observed along the length of the Coorong. These are illustrated in Figure 3.8a. The Murray Mouth region and the North Lagoon existed in the Estuarine/Marine state (characterised by relatively low average salinities, recent freshwater flow over the barrages, high average water depths and high variability of water levels throughout the year). This state supported a diverse fish community, with both estuarine and marine species present, many piscivorous fish and a diverse adult and juvenile benthic invertebrate community. The South Lagoon, however, existed at this time in the Average Hypersaline state (characterised by a less diverse fish fauna, few piscivorous birds, but high numbers of waders and waterfowl). *Ruppia tuberosa*, a key macrophyte for the Coorong, was also widespread, and some salinity-tolerant invertebrate species persisted. For the Average Hypersaline state, the characteristic environmental conditions include higher average salinities than found in the marine basin states, low average depths, but recent freshwater flows over the barrages. Water quality parameters indicated high nutrient concentrations and high phytoplankton abundances.

Pair-wise comparison between the two states suggests that the main differences between the bird faunas were higher abundances of waders (e.g. banded stilt, red-necked avocet, sharp-tailed sandpiper and red-necked stint) and higher waterfowl abundances (e.g. chestnut teal, grey teal, Australian shelduck) in the Average Hypersaline state than the Estuarine/Marine state (see Tables 3.1 and 3.3). The Estuarine/Marine state, on the other hand, supported higher numbers of piscivorous birds (e.g. great cormorant, Caspian tern *Sterna caspia*, Australian white ibis and Australian pelican) than the more hypersaline state. The Estuarine/Marine state also supported higher numbers of adult polychaete worms, juvenile capitellids, and adult *Arthritica helmsi* (a bivalve) while the Average Hypersaline state supported higher numbers of chironomid larvae and a much greater coverage of *Ruppia tuberosa*. The Estuarine/Marine state had a greater abundance and diversity of fish taxa, including mulloway, Australian salmon, bony and black bream, bronze whaler and gummy sharks (*Mustelus antarcticus*) and European carp (*Cyprinus carpio*).

In 2000, the Coorong could have been considered to be in reasonable condition, although it was probably degraded with respect to conditions during the 1980s. While 1997, 1998 and 1999 had relatively small freshwater flows over the barrages, there had been good flows in the years prior to 1997 and, as a result, the Murray Mouth maintained its natural connection to the ocean. For 2000, the Coorong supported a total of 113,500 waterbirds (compared to an average of 153,000 for 2000 to 2007) across 50 species (compared to an average of only 53.25 for 2000 to 2007). The commercial fishery was also reasonably productive (although below average) for that year, with a total of 2.56 kilograms of fish caught per net day (i.e. catch per unit effort, compared to an average of 2.92 for the period of 2000 to 2007) across 14 target species (compared with an average of 13 species for 2000 to 2007).

Along the length of the Coorong, the average water level was 0.37 m AHD at the channel midpoint (as opposed to the thalweg). The average salinity for the system for the year was 48.4 g L^{-1} (compared to approximately 35 g L^{-1} for seawater), with average annual salinity varying along the length of the system from a minimum of 21.3 g L^{-1} near the largest barrage, to 96.8 g L^{-1} at Salt Creek in the South Lagoon. Nutrient concentrations were only measured for the locations south of Pelican Point in 2000, but total phosphate concentrations varied at all those sites between 0.13 and 0.50 mg L^{-1} and total Kjeldahl nitrogen (TKN) concentrations varied between 2.1 and 6.1 mg L^{-1}, with higher readings tending to occur in the South Lagoon.

Figure 3.8. Distribution of ecosystem states along the Coorong

a) In 1999 to 2001, b) In 2002, c) In 2003, d) In 2004 to 2005, e) In 2006, f) In 2007
Note: Dotted lines indicate boundaries between the three regions. Dots indicate the locations of the focal sites (see Figure 1.1 for names). Names of regions are only listed on panel a, but apply to all other panels as well.

It is interesting to note that there is no distinction in the model between conditions near the barrages, with their input of freshwater, and further down the North Lagoon, where this influence is likely to be diminished. This could either be the result of estuarine conditions occurring throughout the northern section of the system, biota occurring in the estuary that was not significantly different from that occurring in the more marine conditions further south, or a biota that varied to the extent that any distinction was not detected statistically. The relatively dry conditions for the three years preceding 2000 (three-year average annual flow of only 2,362 GL per annum compared with the long-term average of 5,050 GL per annum, as calculated from data provided by MDBA) may suggest that the first option is not likely to be the case. This may indicate that conditions in 2000 were not pristine, or that the estuarine biota supported by the system was similar to that supported at somewhat higher salinities.

State of the Coorong in 2007

In addition to varying along the length of the Coorong, the environmental conditions, and the corresponding mix of ecosystem states, have also changed through time. As the drought has continued, and worsened, the more natural states of the Coorong have gradually been replaced by more degraded states. A lack of freshwater has led to a loss of estuarine conditions from the system. With seawater now only source of water replacing evaporative losses, the salt loads within the system have increased over time, with the South Lagoon becoming extremely hypersaline (up to a measured maximum of 220 g L^{-1} in 2007). These changing environmental conditions have resulted in changes in the biological conditions, with the replacement of the original ecosystem states with new ones.

The mix of states observed in 2007 illustrates these changing conditions (Figure 3.8f). Three ecosystem states existed along the length of the Coorong, all of which were different from the ones that were present in 2000. The northern part of the Coorong, including the Murray Mouth region, was predominantly in the Unhealthy Marine state (7 sites), with two sites existing in a Marine state. The Marine state had the lowest average salinity of any state (at 37.4 g L^{-1}), but low average water levels and a long time since flow had passed over the barrages. The water quality of the state reflected the marine inputs through the Murray Mouth, with low values of both nutrient concentrations and turbidity. Biologically, the Marine state supported a relatively diverse marine fish fauna, with some estuarine species (e.g. black bream) also present. It had relatively few piscivorous birds, but many waterfowl and several species of polychaete worms, amongst other invertebrate taxa. Compared with the Marine state, salinity for the Unhealthy Marine state was high, but nutrient concentrations and turbidity remained relatively low. The biota of the Unhealthy Marine state was characterised by high abundances of yellow-eyed mullet in particular, although other marine species were also present in reasonable numbers.

A pair-wise SIMPER comparison between the Marine and Unhealthy Marine states shows that they were distinguished best by the Unhealthy Marine state having twice the abundance of yellow-eyed mullet and small-mouthed hardyhead. The Unhealthy Marine state also had higher abundances of mulloway, European carp, greenback flounder and bony bream, but the Marine state supported more black bream. Differences in bird fauna were most striking in the high abundance of duck species supported by the Unhealthy Marine state (including chestnut teal and Australian shelduck), whereas the Marine state supported higher numbers of Australian white ibis, Eurasian coot (*Fulica atra*) and pied oystercatcher (*Haematopus longirostris*). The benthic invertebrate communities also differed, with the Marine state supporting substantially higher abundances of the polychaetes, *Nephtys australiensis* and *Australonereis ehlersi*, the bivalve *Arthritica helmsi* and the pulmonate gastropod *Salinator fragilis*, while the Unhealthy Marine state had higher numbers of capitellids.

In 2007, the South Lagoon was in the Unhealthy Hypersaline state (Figure 3.8f). This corresponds to higher average salinities than for either of the marine states present in the North Lagoon, and lower average water levels, with little annual variability in water levels.

The water quality was characterised by high nutrient concentrations and high turbidity. For the Unhealthy Hypersaline state, the biotic community was characterised by a low diversity of fish, low numbers of piscivorous birds, but high abundances of banded stilt and Australian shelduck. The invertebrate diversity was low, but chironomid larvae abundances were high.

A pair-wise SIMPER comparison with the Unhealthy Marine state showed that the Unhealthy Hypersaline state supported significantly more banded stilt, hoary-headed grebe, red-necked avocet and teal. On the other hand, the Unhealthy Marine state supported more great and pied cormorants (*Phalacrocorax varius*), crested tern (*Sterna bergii*), curlew sandpiper and Australian pelican. The Unhealthy Marine state had a more diverse invertebrate fauna, supporting more capitellids, *Simplesetia aequisetis*, amphipods and *Arthritica helmsi* than the Unhealthy Hypersaline state. The Unhealthy Marine state supported yellow-eyed mullet, Australian salmon, mulloway, bony and black bream and greenback flounder, all of which were absent from the Unhealthy Hypersaline state.

Like the Unhealthy Marine state, the Marine state supported significantly more fish species than the Unhealthy Hypersaline state, including most of those listed above. The Unhealthy Hypersaline state, on the other hand, supported three times the abundance of small-mouthed hardyhead compared with the Marine state. Banded stilt and Australian shelduck were absent from the Marine state, but present in large numbers in the Unhealthy Hypersaline state, while chestnut and grey teal and red-necked avocet were also more abundant in the latter. Little pied cormorant, Australian white ibis and crested tern were all present in higher numbers in the Marine state, as were the invertebrate taxa *Nephtys australiensis*, *Simplesetia aequisetis*, *Australonereis ehlersi* (all polychaetes) and *Arthritica helmsi*.

In 2007, average water level along the length of system varied between 0.22 m AHD and 0.24 m AHD at its cross-channel midpoint. The average salinity for the system had increased from an average of 38.1 g L^{-1} at the northern end of the system to an average of 170.4 g L^{-1} at Salt Creek, with none of the Coorong sites being below average seawater salinities (overall the average for the system was 77.9 g L^{-1}). Nutrient concentrations had also increased, particularly in the South Lagoon, where the maximum TKN concentration recorded was 10.2 mg L^{-1} and the maximal total phosphate concentration was 1.49 mg L^{-1}, both at Salt Creek.

The biota present in the system had also undergone some dramatic changes. The total number of waterbirds supported in the Coorong remained high at around 153,000, with 55 species represented. However, 42% of these individuals (almost 65,000 birds) were banded stilt, which are more typical of inland salt lakes and feed on brine shrimp. This compared with fewer than 2,500 banded stilt during 2000. Of the remaining population, species of waterfowl like chestnut teal (declined 60%), grey teal (declined 53%) and Australian shelduck (declined 44%) and some iconic species, including fairy tern (declined 56%) and Australian pelican (declined 32%) were amongst the worst-affected. The commercial fishery remained very productive, with 4.71 kilograms per net day caught in the Coorong across 14 species. However, yellow-eyed mullet (*Aldrichetta forsteri*) made up 78% of the total catch, compared with only 40% in 2000, indicating that this species was supporting the vast majority of the local fishing pressure. Changes had also occurred in the benthic macroinvertebrate community. While benthic surveys were not begun until 2004, the South Lagoon then deteriorated from supporting seven taxa at healthy abundances in 2005 to supporting no macroinvertebrate benthos in 2007. Salt-tolerant chironomid larvae were found the furthest south, at Parnka Point (see Figure 1.1).

3.2.   Developing a predictive model

The main task associated with developing a predictive model was in extrapolating data modelled in the hydrodynamic model to the three sites that were outside (or on the border of) the model domain. Table 3.5 presents the equations that were developed for each variable from the ecosystem state model for each of the three sites which needed extrapolated data.

Strong correlations were found between water levels and tidal ranges at Ewe Island and the sites outside the model domain. Only poor relationships were found for salinity, and those presented were the strongest that were identified. Salinity drove the division between Unhealthy Marine and Degraded Marine, which were rarely predicted for any of the northern-most sites, so this was not seen as an issue (see results of scenario analyses).

Variable	Site	Extrapolated from	Equation	Sum of Squares
Simple regressions				
Average water level	Monument Road	Ewe Island	y = 1.008 x Average water level	0.801
	Mundoo Channel	Ewe Island	y = 0.939 x Average water level	0.648
	Barkers Knoll	Ewe Island	y = 1.413 x Average water level	0.878
Average tidal range	Monument Road	Ewe Island	y = 1.714 x Average tidal range	0.712
	Mundoo Channel	Ewe Island	y = 1.014 x Average tidal range	0.986
	Barkers Knoll	Ewe Island	y = 0.923 x Average tidal range	0.990
Multiple regressions				
Average salinity	Monument Road	Ewe Island	y = 1.261 x Average salinity - 0.000045 x (Average monthly flow + 1)	0.495
	Mundoo Channel	Ewe Island	y = 1.156 x Average salinity - 0.000029 x (Average monthly flow (month - 1) + 1)	0.231
	Barkers Knoll	Ewe Island	y = 1.240758 Average salinity - 0.000015 x (Average monthly flow + 1)	0.145

Table 3.5. Equations developed to extrapolate modelled data from Ewe Island and Pelican Point to the sites outside the domain of the hydrodynamic model for the alternative ecosystem state model

Note that other relevant variables in the models (e.g. water level from the previous year) were calculated from these values. Unless specified, all averages are annual averages.

## 3.3. Alternative models independent of barrage flows

One of the key driving parameters for the ecosystem model described above was the occurrence of freshwater flows over the barrages. This meant that only limited changes in ecological conditions could be modelled unless such flows were present. Given several of the scenarios to be investigated were designed to be alternatives to having freshwater flows in the short term, we developed an alternative set of new models to describe the behaviour of the system without reference to the flows over the barrages.

As mentioned above, the best results for the alternative model was produced when the marine and hypersaline basins were modelled independently. The model for the marine basin (assumed to occur in the North Lagoon under the current conditions) is shown in Figure 3.9. It describes the ecosystem state of the Coorong relative to the water level, the previous year's water level and depth from two years ago. This model correctly classified 72% of the training data set used and 70% of the test data set, indicating that it discriminated well between the marine ecosystem states.

The hypersaline basin model (used to describe current South Lagoon states) identified a combination of average water level, water level from the previous year, the range in water levels over the year (i.e. change between the maximum and minimum water level over the year) and the maximum salinity for the year as driving the ecosystem state of the basin (Figure 3.10). The hypersaline basin model correctly classified 87% of the training data set and 80% of the test data set under cross-validation. This is quite a high degree of predictive success given the variability inherent in ecological data sets.
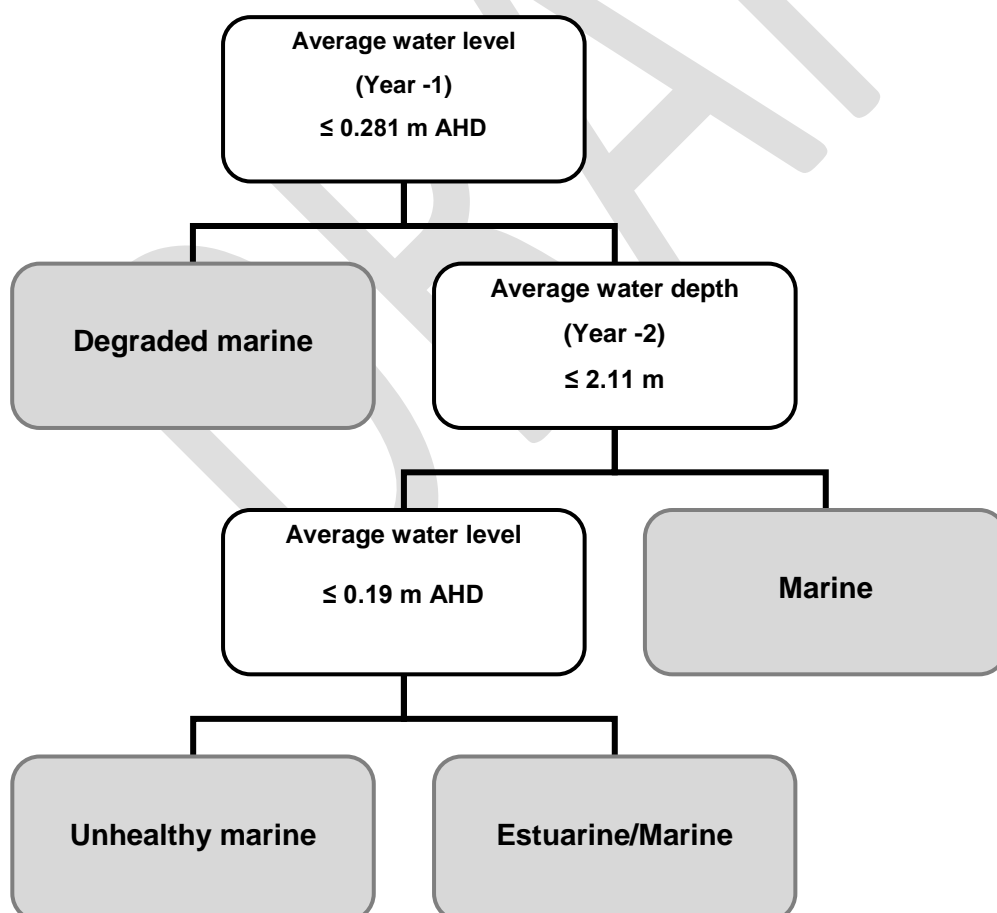

Figure 3.9 Marine (or northern) basin model for the Coorong excluding flow parameters as predictive variables.
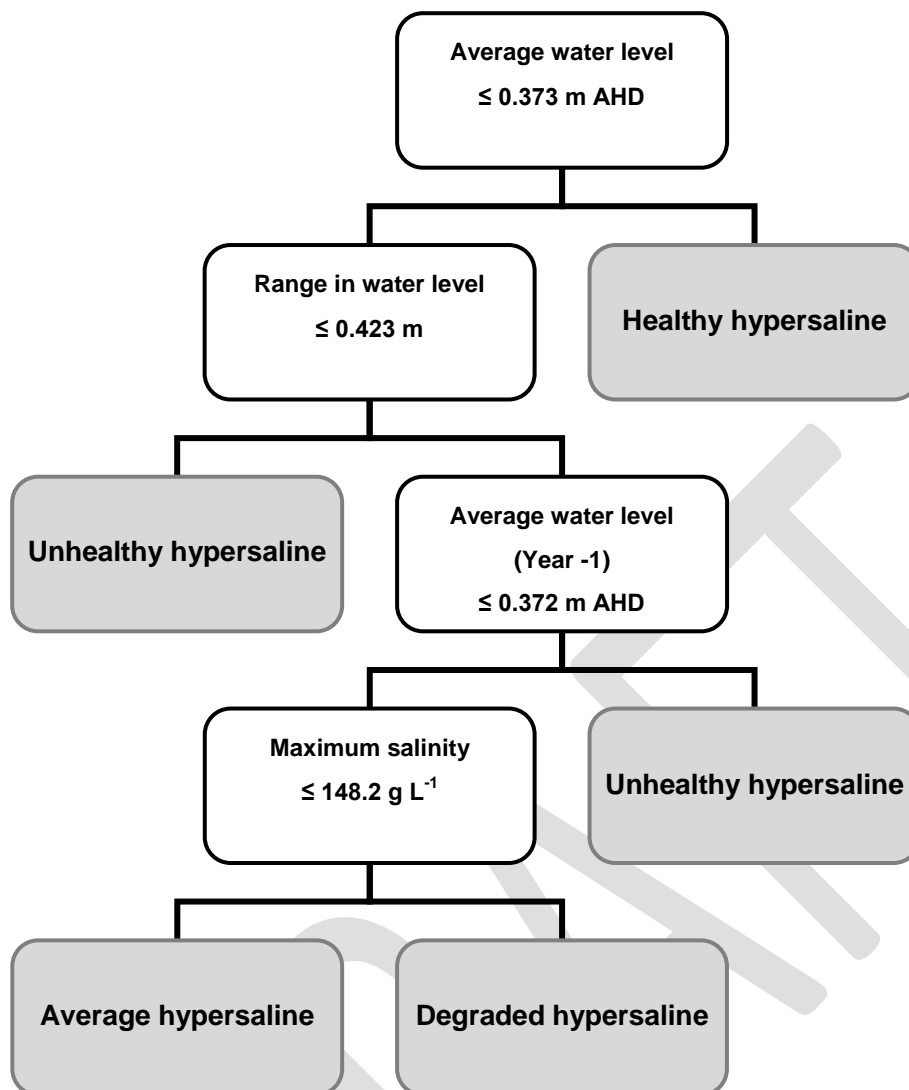
Figure 3.10 Hypersaline (southern) basin model for the Coorong excluding flow parameters as predictive variables.

Note: The Unhealthy Hypersaline state appears in the model twice, indicating that there are two distinct pathways to reach that state.

When the Baseline scenario was assessed using both the original and alternative models, the biggest differences was in the proportion of site-years predicted to be in the Marine state (Figure 3.11). For the original model, this was small (1%) while for the alternative model it was much higher (24%). However, the proportion of site-years predicted to be in either the Estuarine/Marine state or the Marine state was the same for both models (72%). This difference also affected the proportion of site-years predicted to be in a degraded state (32% for the alternative model versus 6% in the original model) as the Marine state is considered degraded by the definition used (relating to the length of time since barrage flows). The proportion of site-years predicted to be in the remaining states was quite consistent across the two models. The other noticeable discrepancy was that the alternative model tended to predict more site-years as being in the Degraded Hypersaline state (3% versus less than 1% in the original model). Overall the concordance in the predictions for individual site-years was relatively modest at 56%.
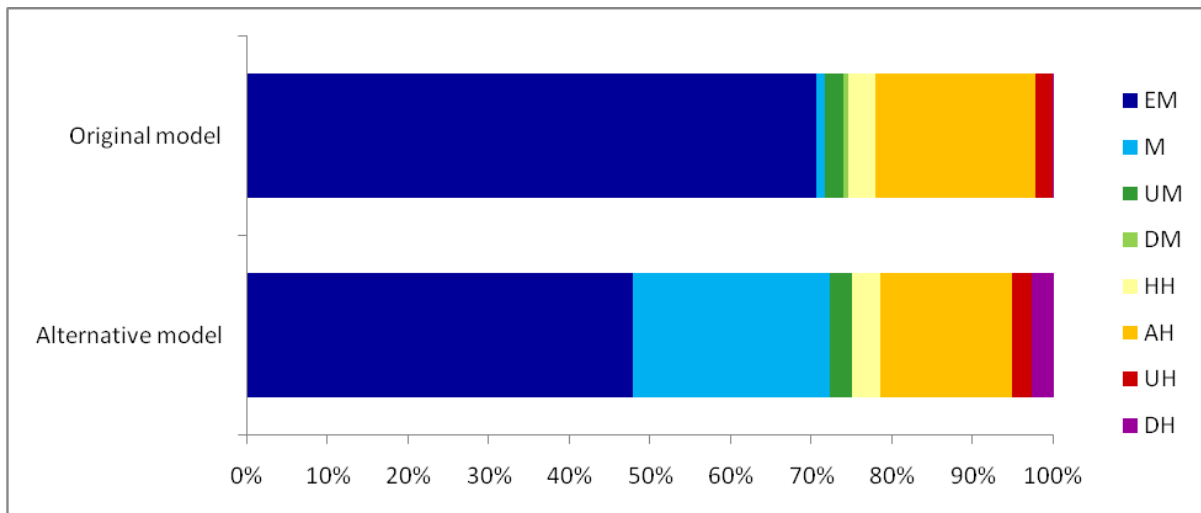
Figure 3.11. Proportion of site-years predicted to be in each state for the Baseline scenario assessed with the original ecosystem state model compared with the alternative ecosystem state model over 114 years for 12 focal sites.

Note: EM = Estuarine/Marine, M = Marine, UM = Unhealthy Marine, DM = Degraded Marine, HH = Healthy Hypersaline, AH = Average Hypersaline, UH = Unhealthy Hypersaline, DM = Degraded Hypersaline.

As for the original model, variables for three northernmost sites were extrapolated from the modelled data for Ewe Island (Figure 1.1). Table 3.6 gives the equations that were developed for each of the variables for each site.

Water levels and depths were able to be extrapolated with a high degree of accuracy, but maximum salinities and ranges in water levels had poor relationships. Fortunately, both of these variables were part of the hypersaline model (rather than the marine model) which was not applied to the three northern-most sites in the Coorong, so these poor relationships did not affect the predicted ecosystem states under the alternative model.

Variable	Site	Extrapolated from	Equation	R^2
Simple regressions				
Average water level	Monument Road	Ewe Island	y = 1.001 x Average water level	0.801
	Mundoo Channel	Pelican Point	y = 0.894 x Average water level	0.502
	Barkers Knoll	Pelican Point	y = 1.414 x Average water level	0.878
Multiple regressions				
Range in water level	Monument Road	Ewe Island	y = 0.891 x Range in water level + (0.713 x Maximum tidal range)	0.337
	Mundoo Channel	Ewe Island	y = 0.588 x Range in water level + (1.006 x Maximum tidal range)	0.271
	Barkers Knoll	Ewe Island	y = 0.490 x Range in water level + (1.009 x Maximum tidal range)	0.976
Average water depth (Year -2)	Monument Road	Ewe Island	y = 0.782 x Average depth (Year -2) + (1.366 x Average depth)	0.933
	Mundoo Channel	Ewe Island	y = 0.844 x Average depth (Year -2) + (0.461 x Average depth)	0.993
	Barkers Knoll	Ewe Island	y = 1.923 x Average depth (Year -2) + (0.830 x Average depth)	0.715
Maximum salinity	Monument Road	Ewe Island	y = 1.261 x Maximum salinity 1 (0.000045 x Average flow)	0.287

Table 3.6. Equations developed to extrapolate modelled data from Ewe Island and Pelican Point to the sites outside the domain of the hydrodynamic model for the alternative ecosystem state model

Note that no useful relationships were identified for maximum salinity for either Mundoo Channel or Barkers Knoll, but as maximum salinity was a variable from the hypersaline basin model (see Figure 3.10) while both sites were allocated to the marine basin, this did not pose a problem for predicting ecosystem states at these sites. Other relevant variables in the models (e.g. water level from the previous year) were derived from these values (e.g. water level from the previous year is water level with one-year lag, so a separate regression is not required).

3.4. Model evaluation

3.4.1. Comparing model predictions for 1980s to available data

An attempt was made to evaluate the model predictions by comparing the biota from the predicted ecosystem state to the available data for the 1980s and 1990s. When the available data for individual taxonomic groups from the 1980s and 1990s were compared with the equivalent data for the years within the training data set (1999-2007), results were mixed.

For commercial CPUE for fish species, there was a significant interaction between states and years, with significant differences between states ($Pseudo\text{-}F$ = 5.2397, p = 0.001; Figure 3.12), but also significant differences between years ($Pseudo\text{-}F$ = 1.986, p = 0.001, with years nested within states). Contrasts between states from the 1980s and 1999-2007 and between the 1990s and 1999-2007 were also significant ($Pseudo\text{-}F$ = 6.2731, p = 0.001 for the 1980s and $Pseudo\text{-}F$ = 2.5275, p = 0.012 for the 1990s) indicating that the same state showed different in biotic composition between decades. The dispersion within states also varied significantly between states (F = 4.2206, p = 0.019).
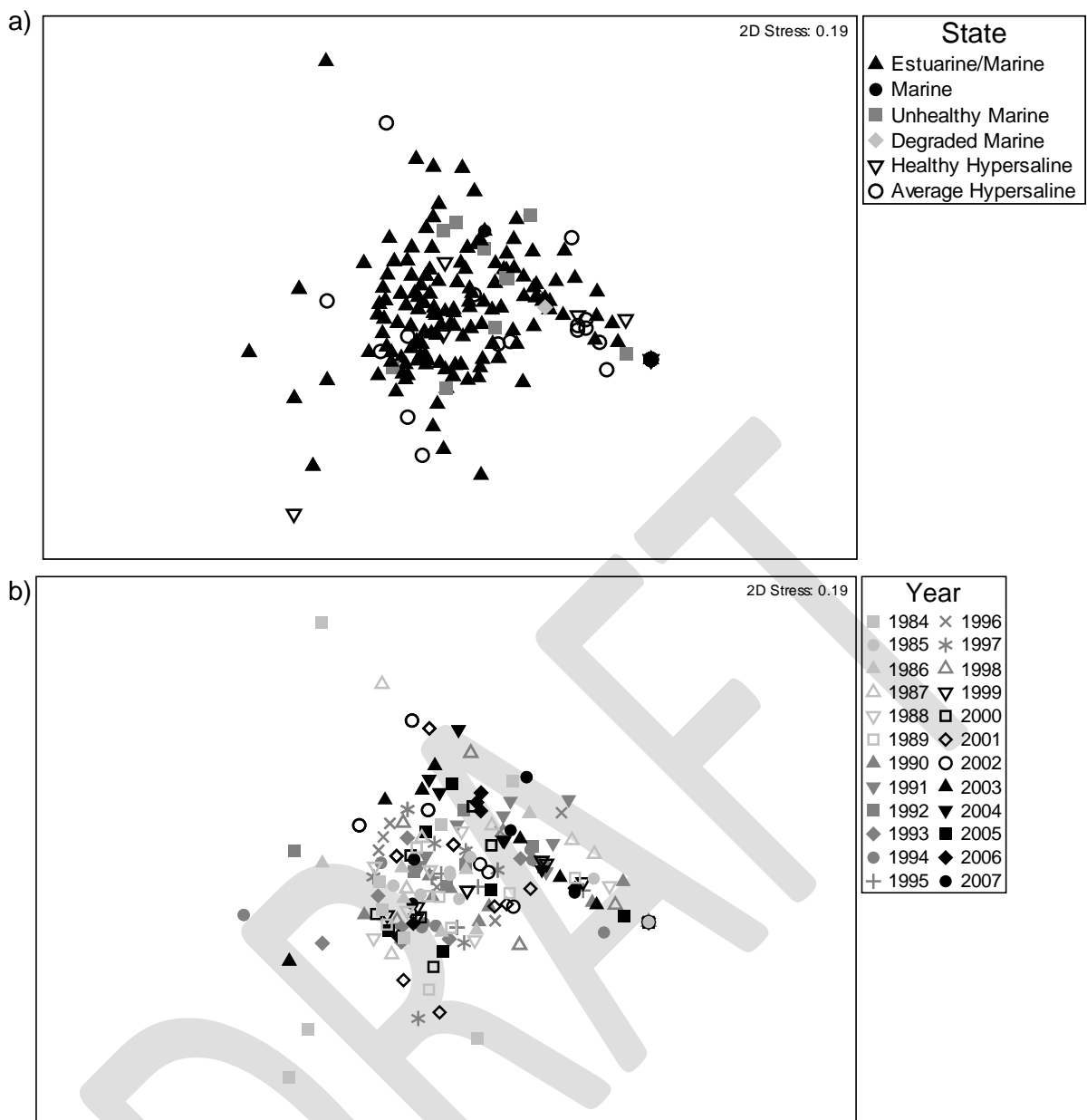
Figure 3.12. nMDS plot of catch per unit effort for commercial fish species between 1984 and 2007

a) Coded to show differences by state, b) Coded to show differences by year

Note: The figure is based on a Bray-Curtis similarity matrix of standardised, log(x+1)-transformed abundance data for commercial fish CPUE for 12 sites between 1984 and 2007 (minus cases with missing values, $n = 250$). A maximum of 25 runs were used.

Macroinvertebrate assemblages in 1982 at Parnka Point were distinctly different from those at remainder of the sites across the rest of the years sampled (Figure 3.13a). These differences drove significant differences between the states (*Pseudo-F* = 3.4844, $p = 0.018$, Figure 3.13b), and across the years (*Pseudo-F* = 5.5604, $p = 0.001$, Figure 3.13c). The invertebrate assemblages from the 1980s were significantly different from those in the 2000s (*Pseudo-F* = 14.034, $p = 0.001$). Again, there were significant differences in the level of dispersion within states ($F = 12.754$, $p = 0.003$).
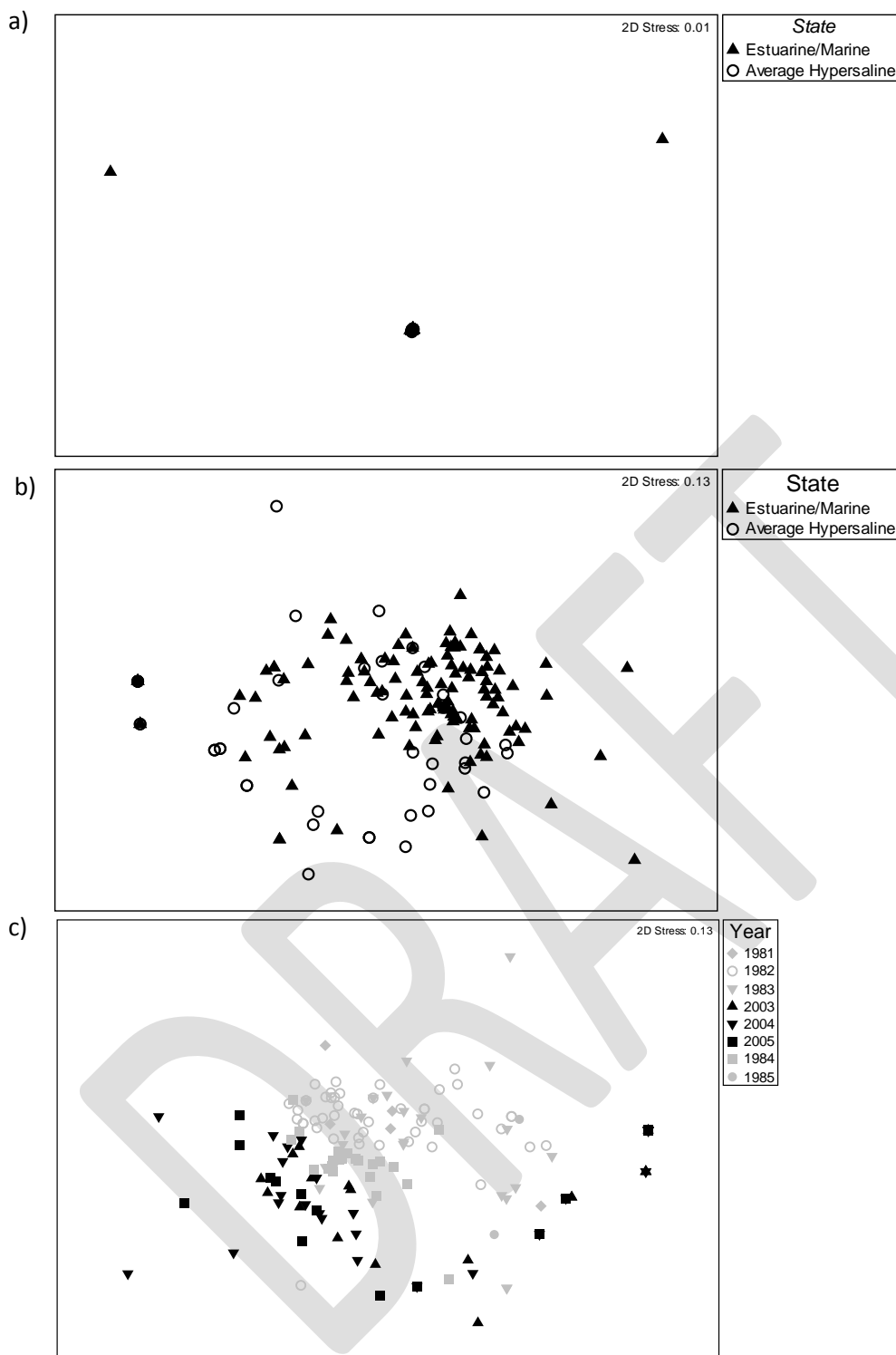
Figure 3.13. nMDS plot based on the presence/absence of invertebrate species for 1981 to 1985 and 2003 to 2005

a) Whole plot, b) Subset of centre of plot shown in a (*n* = 169), coded by state, c) Subset of plot shown in a (*n* = 169), coded by year (both excluded points were from 1982)

Note: The figure is based on a Bray-Curtis similarity matrix of presence/absence data for invertebrate species for seven sites in 1981 to 1985 and 2003 to 2005 (minus cases with missing values, *n* = 169). A maximum of 25 runs were used.

Two comparisons were possible from the available bird data. The first involved a limited number of species over four regions in the Coorong from 1982, 1984-85, 1987 and 2000-2007 (Figure 3.14). Differences between states were significant (*Pseudo-F* = 2.3734, *p* = 0.039, Figure 3.14a), but differences between years were not (*Pseudo-F* = 0.71881, *p* = 0.899, Figure 3.14b). A direct comparison between states occurring in the 1980s and those in the 2000s was also not significant (*Pseudo-F* = 1.5376, *p* = 0.197), nor were differences in dispersion within states found.
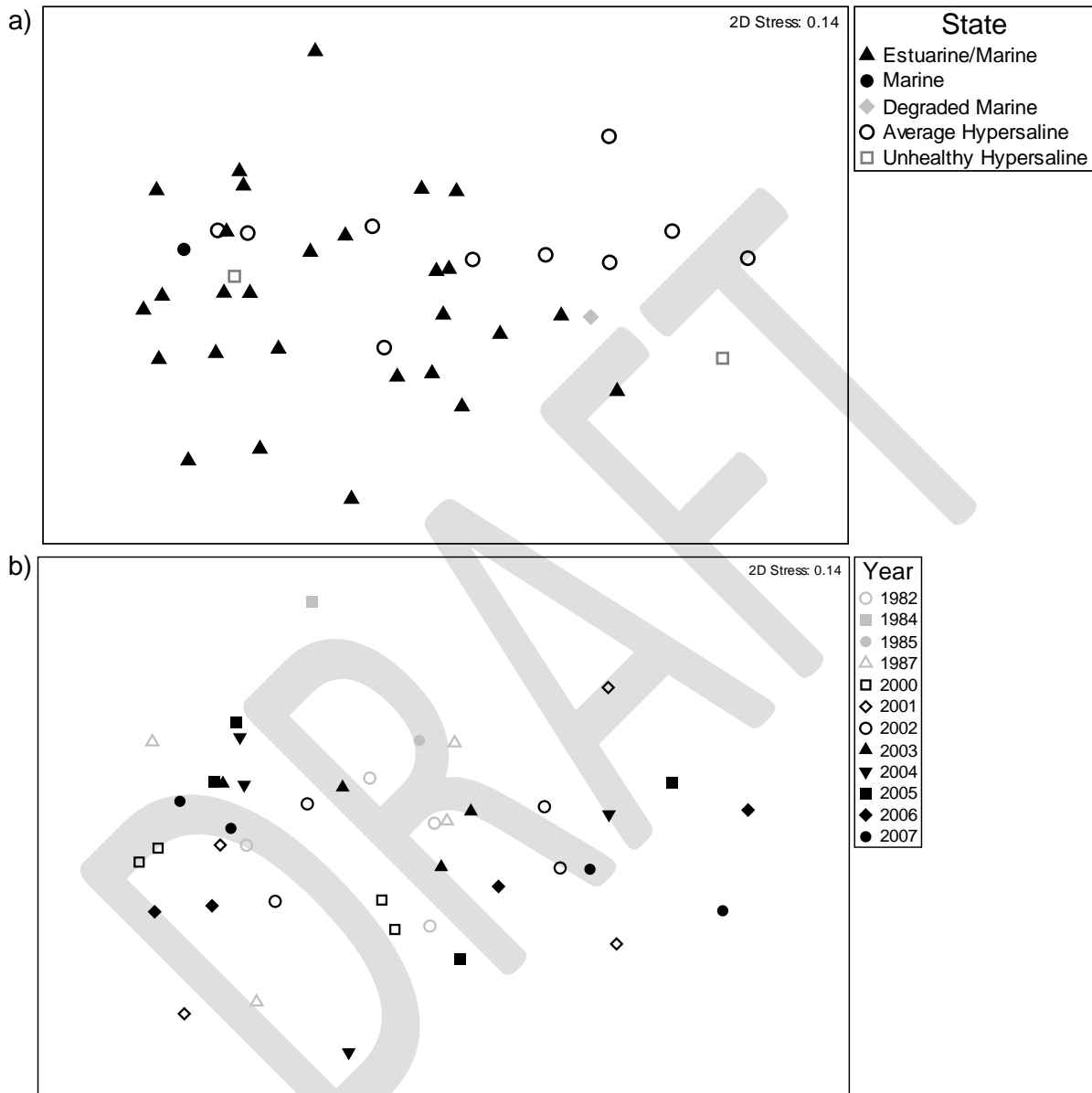
Figure 3.14. nMDS plot based on bird abundance for particular species for 1982, 1984, 1985, 1987 and 2000 to 2007

a) Coded to show differences by state, b) Coded to show differences by year

Note: The figure is based on a Bray-Curtis similarity matrix of standardised, log(x+1)-transformed abundance data for a limited number of bird species in the years mentioned over four regions (Murray Mouth, North Lagoon, around Parnka Point, South Lagoon) in the Coorong (minus cases with missing values, *n* = 42). A maximum of 25 runs were used.

A second analysis was done using a wider range of bird species (all that were present) for a smaller number of years (1984-1985, 2000-2007). The nMDS for this data (Figure 3.15) seemed to show better grouping by state than any of the previous analyses (Figures 3.12 to

3.14), however, neither states (*Pseudo-F* = 1.5751, *p* = 0.144, Figure 3.15a) nor years (*Pseudo-F* = 0.60276, *p* = 0.98, Figure 3.15b) proved significantly different. A comparison between the states observed in the 1980s and those in the 2000s, did however reveal significant differences (*Pseudo-F* = 2.8873, *p* = 0.031) and there were significant differences in the level of dispersion within states (*F* = 7.2408, *p* = 0.024).

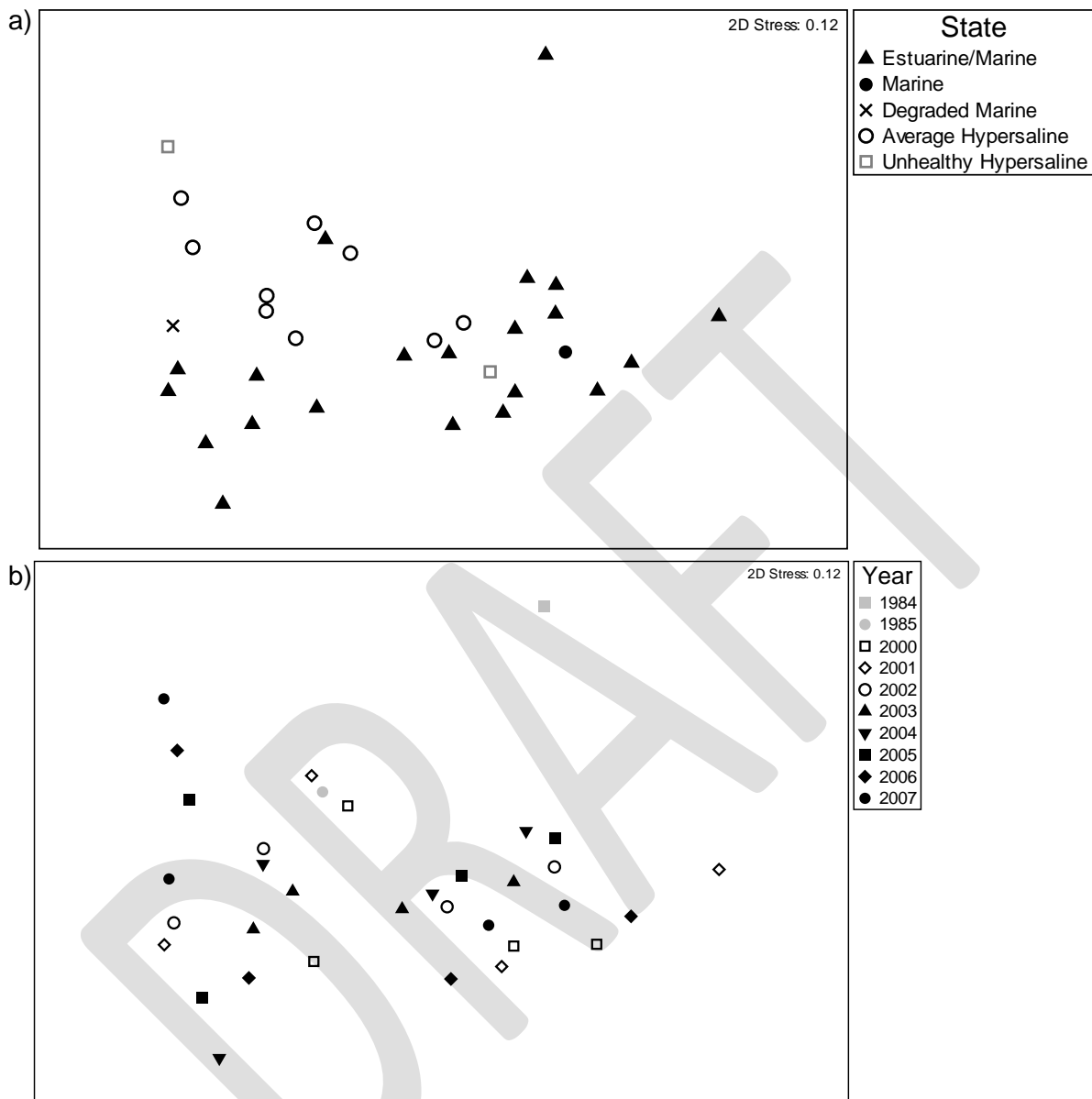**Figure 3.15. nMDS plot based on bird abundance for 1984, 1985 and 2000 to 2007**

a) Coded to show differences by state, b) Coded to show differences by year

Note: The figure is based on a Bray-Curtis similarity matrix of standardised, log(x+1)-transformed abundance data for bird species in the years mentioned over four regions (Murray Mouth, North Lagoon, around Parnka Point, South Lagoon) in the Coorong (minus cases with missing values, *n* = 42). A maximum of 25 runs were used.

Ruppia tuberosa shoot and turion coverage in the Coorong was compared for 1985, 1990-1993 and 2000-2007 (Figure 3.16). This analysis did not detect significant differences due to ecosystem state (*Pseudo-F* = 0.80401, *p* = 0.542, Figure 3.16a), nor by year (*Pseudo-F* = 0.94206, *p* = 0.561, Figure 3.16b). Neither the ecosystem states in the 1980s nor those in the 1990s differed from their counterparts in the 2000s (*Pseudo-F* = 2.6274, *p* = 0.120 for the 1980s and *Pseudo-F* = 2.0657, *p* = 0.190 for the 1990s). There was, however, a significant difference in the dispersion within ecosystem states (*F* = 3.1443, *p* = 0.039).
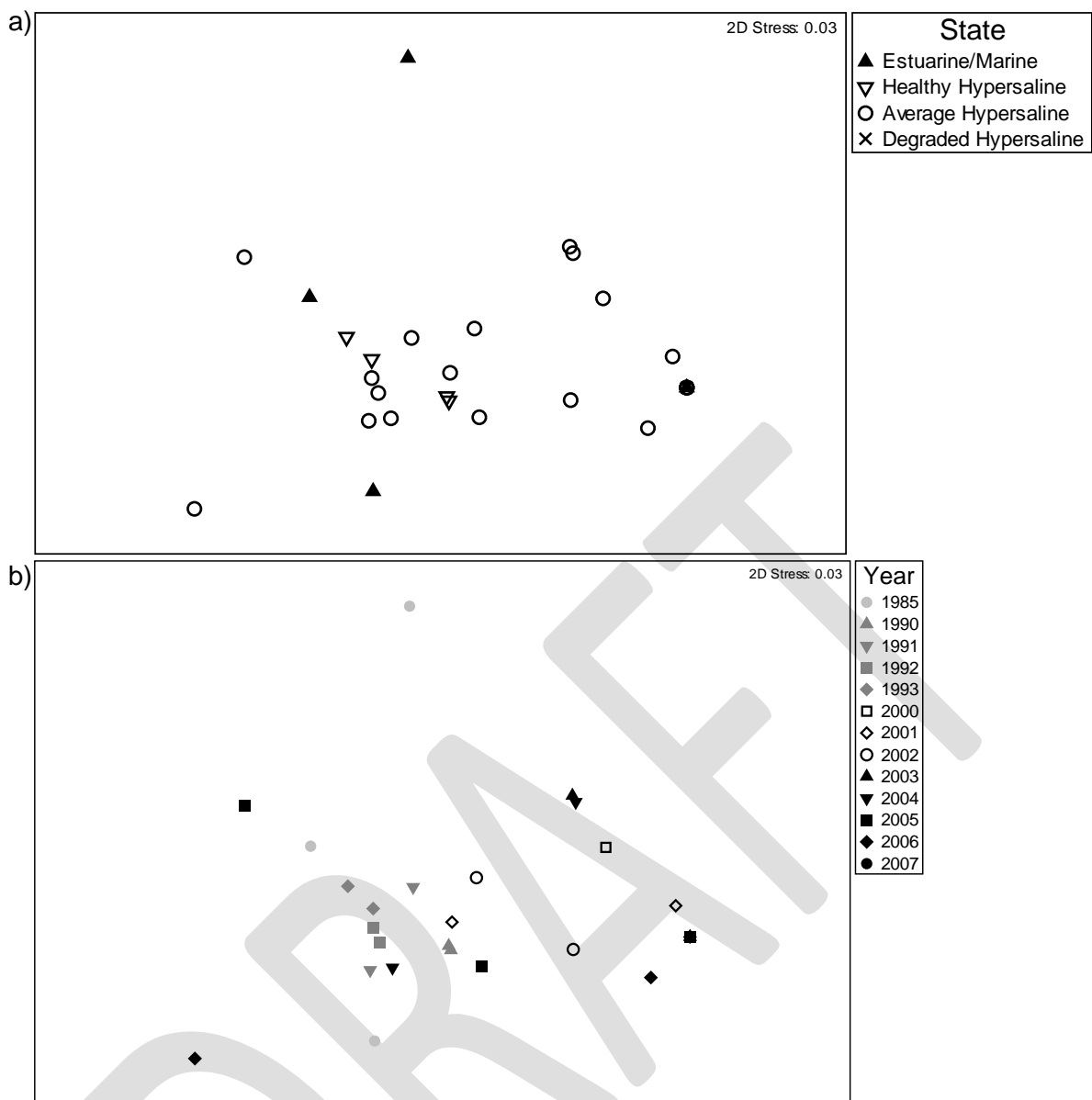
Figure 3.16. nMDS plot based on Ruppia tuberosa *shoot and propagule coverage for 1985, 1990-1993 and 2000 to 2007*

a) Coded to show differences by state, b) Coded to show differences by year

Note: The figure is based on a Bray-Curtis similarity matrix of standardised, log (x+1)-transformed coverage data for *Ruppia tuberosa* shoots and propagules at five sites in the Coorong (minus cases with missing values, *n* = 45). A maximum of 25 runs were used.

In summary, based on these analyses, the commercial fish, invertebrate and the bird analysis based on a limited number of species suggested that the biota found in each ecosystem state was significantly different. In contrast, the analysis based on a more extensive list of bird species (but for fewer years) and on *Ruppia tuberosa* did not detect differences between ecosystem states, perhaps suggesting that birds and macrophytes occurred more extensively (or more consistently) throughout the Coorong.

Analyses based on the commercial fish, invertebrates and the more extensive bird abundances indicated that states from previous decades were biologically distinct from their counterparts in the training period for the ecosystem state model (1999-2007). This was not the case for the more limited bird abundance analysis nor that based on *Ruppia tuberosa*.

These results are somewhat contradictory given that there are differences between states and between years, or significant interactions between the two for some taxa, but not others. Thus,

it is difficult to gain a clear understanding of whether a) the ecosystem state support distinct biota across the range of taxonomic groups, and b) whether there is evidence that the biota from the 1980s (and the 1990s) differed from that found during the 1999-2007 period on which this model is based. To resolve these issues, we attempted several multi-group analyses for the Coorong. A summary of all single-taxonomic group comparisons are given in Table 3.7, showing the data included, the sample size, comparisons made and the outcomes of each comparison.

Of the analyses involving multiple taxonomic groups, the most comprehensive in terms of the number of groups included the presence/absence of aquatic macrophytes, commercial fish species and invertebrates. This also had the fewest cases ($n = 8$). No significant differences was detected between the two ecosystem states present (i.e Estuarine/Marine and Average Hypersaline) in the data set ($Pseudo\text{-}F = 2.1971$, $p = 0.099$, Figure 3.17a), but ecosystem states from the 1980s did differ from those in the 2000s ($Pseudo\text{-}F = 4.9393$, $p = 0.044$; Figure 3.17b). The dispersion of cases for the two states were significantly different ($F = 22.593$, $p = 0.023$).

The average similarity between cases in the Estuarine/Marine state was only 50.80%. The top three species contributing to that similarity were mulloway, yellow-eye mullet and chironomid sp. No other macroinvertebrate species were identified as typifying the state. For the Average Hypersaline state, the average similarity between cases was 91.23%. Eleven fish species and thirteen invertebrate species were identified as all contributing slightly (and equally) to the similarity between cases.

The average dissimilarity between the Estuarine/Marine and Average Hypersaline states was 44.61%. The Average Hypersaline state tended to have a higher likelihood that most species would be present, including bronze-whaler sharks, *Amarinus lacustris* (a species of crab) and *Phyllodoce novohollandiae* (a polychaete). The Estuarine/Marine state, however, was more likely to have the dipteran *Ephydrella* sp. and *Capitella* sp. present.

Taxa included	Sample Size (n)	Factors	Outcome	Figure Number
Commercial Fish	250	Years: 1984-89; 1990-1998; 1999-2007 States Dispersion	Significant interaction between states & years ($p = 0.001$)	Figure 3.12a
			Significant difference between years nested in states ($p = 0.001$)	Figure 3.12b
			Significant contrast 80s vs 1999-2007 ($p = 0.001$)	
			Significant contrast 90s vs 1999-2007 ($p = 0.012$)	
			Significant dispersion within states ($p = 0.019$)	
Macroinvertebrates	169	Years: 1981-1985; 2003-2005 States Dispersion	Significant difference for 1982, Parnka Point from rest sites & years	Figure 3.13a
			Significant difference between states ($p = 0.018$)	Figure 3.13b
			Significant difference across years nested in states ($p = 0.001$)	Figure 3.13c
			Significant contrast 80s vs 2000s ($p = 0.001$)	
			Significant dispersion within states ($p = 0.003$)	
Birds	42	Years: 1982; 1984-85; 1987; 2000-2007 States Dispersion	Significant difference between states ($p = 0.039$)	Figure 3.14a
			No significant difference between years nested in states ($p = 0.899$)	Figure 3.14b

Table 3.7. Summary of comparisons made between biotic compositions of predicted ecosystem states in the 1980s and 1990s to those of the training data set (1999-2007)

Taxa included	Sample Size (*n*)	Factors	Outcome	Figure Number
Birds	42	Years: 1982; 1984-85; 1987; 2000-2007 States Dispersion	No significant contrast between 80s & 2000s states ($p = 0.197$) No significant dispersion within states	
Birds	42	Years: 1984-85; 2000-2007 States Dispersion	No significant difference between states ($p = 0.144$) No significant difference between years nested in states ($p = 0.98$) Significant contrast between 80s and 2000s states ($p = 0.031$) Significant dispersion within states ($p = 0.024$)	Figure 3.15a Figure 3.15b
Macrophytes (*R.tuberosa* **shoots & turions**)	45	Years: 1985; 1990-93; 2000-2007 States Dispersion	No significant difference between states ($p = 0.542$) No significant difference across years nested in states ($p = 0.561$) No significant contrast 80s vs 2000s ($p = 0.120$) No significant contrast 90s vs 2000s ($p = 0.190$) Significant dispersion within states ($p = 0.039$)	Figure 3.16a Figure 3.16b

Table 3.7 cont. Summary of comparisons made between biotic compositions of predicted ecosystem states in the 1980s and 1990s to those of the training data set (1999-2007)
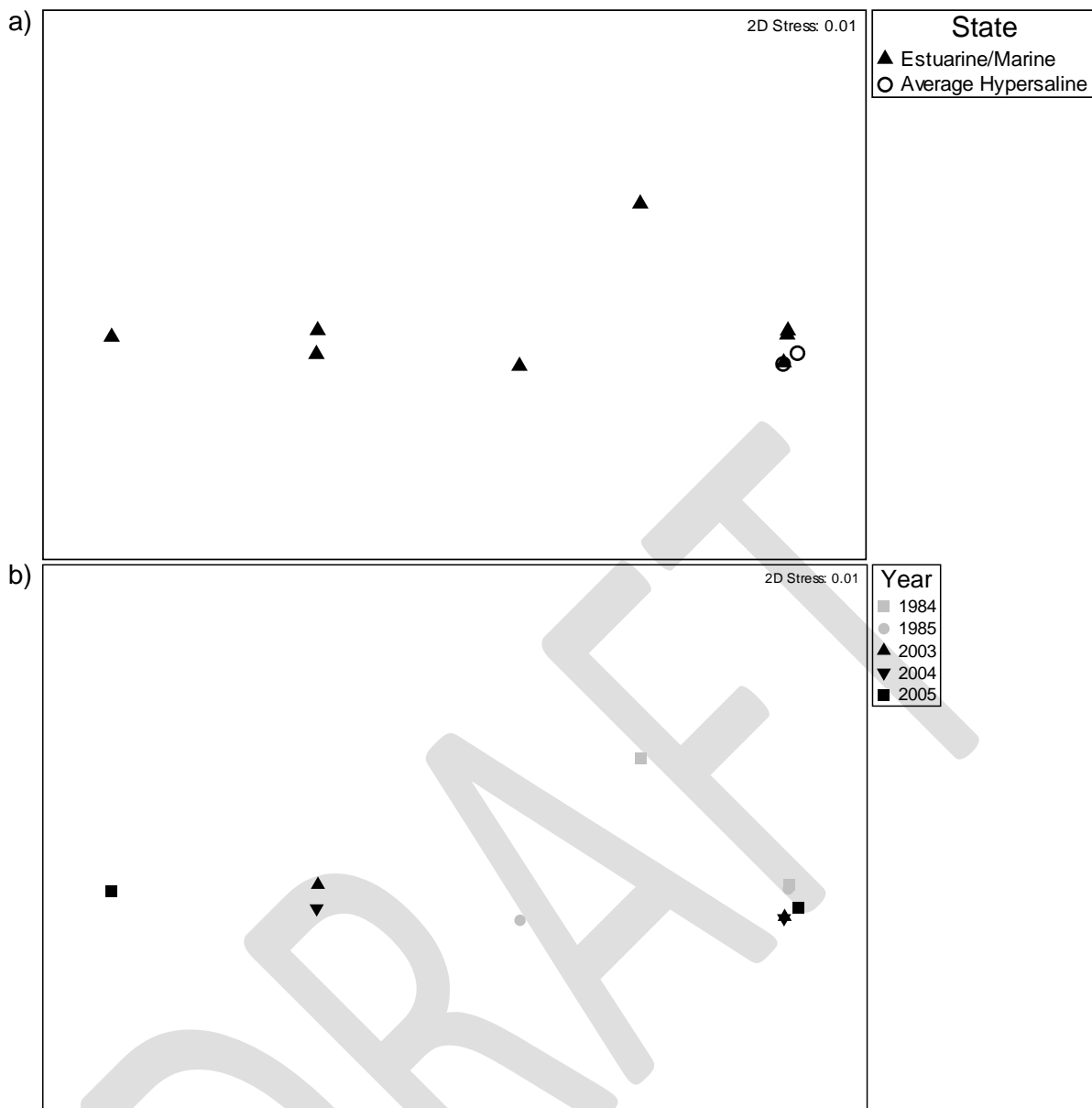
Figure 3.17. nMDS plot based on presence/absence data for aquatic macrophytes, commercial fish and invertebrate species for 1984, 1985 and 2003 to 2005

a) Coded to show differences by state, b) Coded to show differences by year

Note: The figure is based on a Bray-Curtis similarity matrix of presence/absence data for aquatic macrophytes, commercial fish species and invertebrate species for four regions (Murray Mouth, North Lagoon, around Parnka Point, South Lagoon) in the Coorong for the years mentioned (minus cases with missing values, $n = 8$). A maximum of 25 runs were used.

When aquatic macrophytes and commercial fish species presence/absence data ($n = 30$) were analysed, significant differences were identified in the biotic assemblages between states (*Pseudo-F* = 7.0083, $p = 0.001$; Figure 3.18a). No differences were detected between years, nor between states comparing those found in the 1980s with those in the 2000s (*Pseudo-F* = 0.671, $p = 0.825$ and *Pseudo-F* = 2.0994, $p = 0.120$, respectively, Figure 3.18b). PERMDISP indicated that there were significant differences in the amount of dispersion within states ($F = 8.1406$, $p = 0.001$).

Fish species contributed the bulk of the similarity between cases in the Estuarine/Marine state. Average similarity for this state was 78.13%, with Australian salmon, black bream and flounder the top three species typifying the state. The Average Hypersaline state had an average similarity of 50.35%. *Ruppia tuberosa* contributed 47% of the similarity between cases, with

yellow-eyed mullet, mulloway and black bream also typifying the state. No species could be identified as typical for the Marine, Healthy Hypersaline and Unhealthy Hypersaline as each was represented by a single case.

The Estuarine/Marine state and the Average Hypersaline state had an average dissimilarity of 63.49%. The Estuarine/Marine state was more likely to support Australian salmon, black bream, flounder and mulloway, while the Average Hypersaline state was more likely to support *Ruppia tuberosa*. Species contributing to the dissimilarity between the other states have not been presented due to the presence of only a single case for the remaining states.

Figure 3.18. nMDS plot based on presence/absence data for aquatic macrophytes and commercial fish species for 1984, 1985 and 2000 to 2006

a) Coded to show differences by state, b) Coded to show differences by year
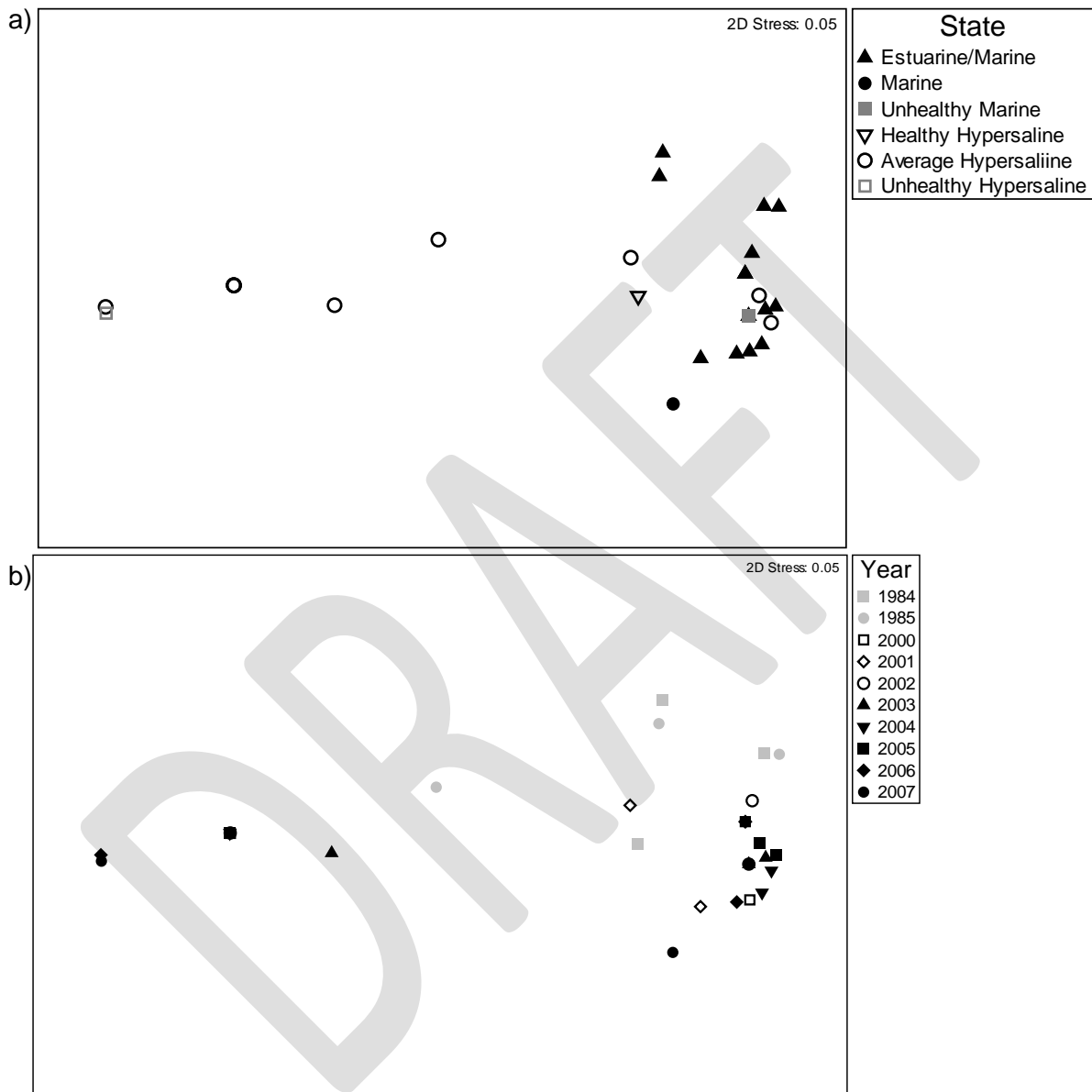
Note: The figure is based on a Bray-Curtis similarity matrix of presence/absence data for aquatic macrophytes and commercial fish species for four regions (Murray Mouth, North Lagoon, around Parnka Point, South Lagoon) in the Coorong for the years mentioned (minus cases with missing values, n = 30). A maximum of 25 runs were used.

The final multi-taxonomic-group comparison investigated differences between commercial fish species and the limited number of bird species presence or absence (Figure 3.19). No significant difference was detected between states (*Pseudo-F* = 1.8865, p = 0.085, Figure

3.19a, b) or years (*Pseudo-F* = 0.41629, *p* = 0.988, Figure 3.19c), but there was a significant interaction between states and years, with 1980s states being significantly different from those in the 2000s (*Pseudo-F* = 27.474, *p* = 0.001).

The average similarity within the Estuarine/Marine state was 78.47%, with the top three species typifying those cases being mulloway, yellow-eye mullet and black-winged stilts. Other species with the same contribution to similarity included curlew sandpiper, common greenshank, red-capped plover, red-necked stint and sharp-tailed sandpiper. The average similarity within the Average Hypersaline state was 78.23%. Here the top species typifying the state included banded stilt, curlew sandpiper, common greenshank, pied oystercatcher, red-capped plover, red-necked avocet, red-necked stint and sharp-tailed sandpiper. Other states had fewer than two cases, so species contributing to similarity could not be calculated.

The average dissimilarity between the Estuarine/Marine and the Average Hypersaline states was 23.40%. The Estuarine/Marine state was more likely to support fish species like black bream, flounder and Australian salmon, while the Average Hypersaline species supported masked lapwing, great-crested grebe and black-faced cormorant more frequently.
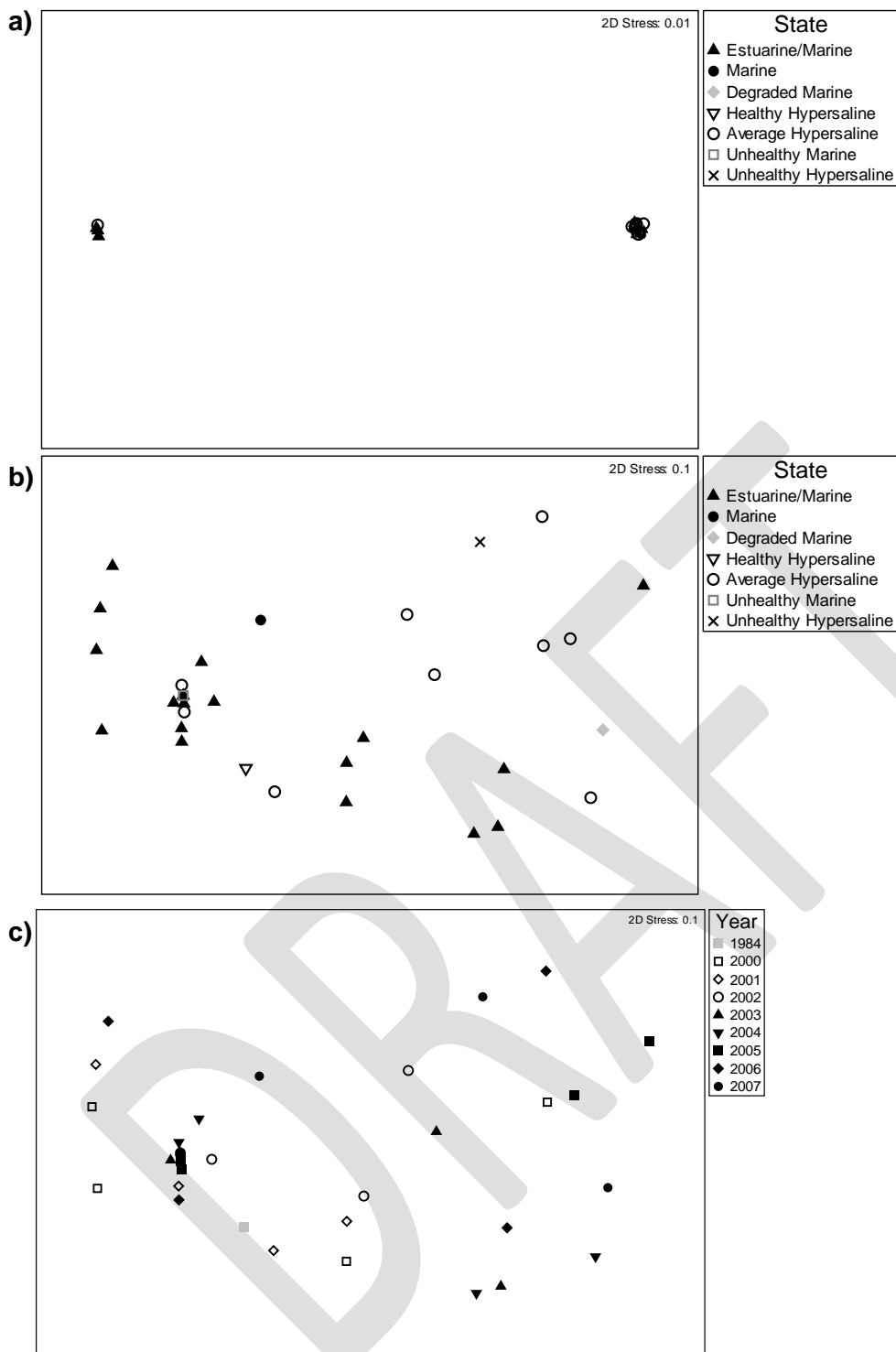
Figure 3.19. nMDS plot based on presence/absence data for commercial fish species and a limited number of bird species for 1984, 1987 and 2000 to 2006

a) Whole plot, b) Subset of centre of plot shown in a (*n* = 34), coded by state, c) Subset of plot shown in a (*n* = 34), coded by year (all three excluded points were from 1987)

Note: The figure is based on a Bray-Curtis similarity matrix of presence/absence data for commercial fish and a limited number of bird species for four regions in the Coorong for the years mentioned (minus cases with missing values, *n* = 37). A maximum of 25 runs were used.

### 3.4.2. Comparing model predictions for 2008 to available data

The only data set that was available for 2008 suitable for comparison with the training data set was fisheries-independent fish data collected across 11 sites. These data were analysed quarterly due to the limited number of years available (Figure 3.20). PERMANOVA indicated that there was no significant difference between the fish biota occurring across states (*Pseudo-F* = 3.386, *p* = 0.130, Figure 3.20a, b), but that years were significantly different (*Pseudo-F* = 2.006, *p* = 0.008, Figure 3.20c) and there was a significant interaction between state and years, with fish assemblages per state in 2008 differed significantly from those in 2006-2007 (*Pseudo-F* = 2.006, *p* = 0.008). There was also a significant difference in the level of dispersion found within states (*F* = 7.2097, *p* = 0.001).
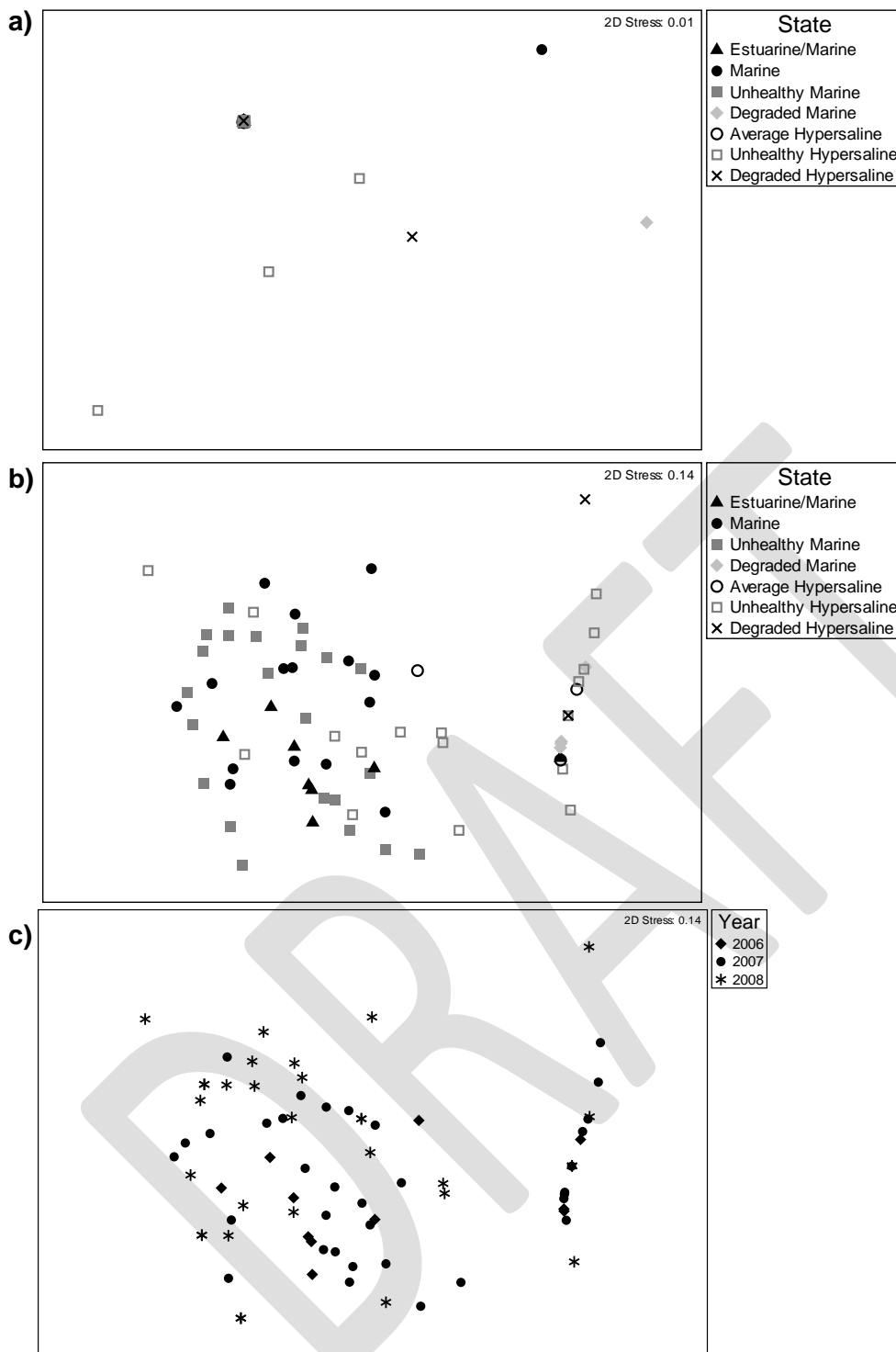
Figure 3.20. nMDS plot based on the fisheries-independent abundance of fish species for 2006 to 2008 based on a quarterly time-step

a) Whole plot, b) Subset of centre of plot shown in a (*n* = 76), coded by state, c) Subset of plot shown in a (*n* = 76), coded by year (two of the six excluded points were from 2008 (the isolated Marine and Degraded Hypersaline points), the others were from 2007)

Note: The figure is based on a Bray-Curtis similarity matrix of log-transformed(x+1) fish abundance data for 11 sites in 2006 to 2008 (minus cases with missing values, *n* = 82). A maximum of 25 runs were used.

3.5. Evidence for any additional ecosystem states within the Coorong

Cluster analysis indicated that the Estuarine/Marine state could be divided into at least three significant clusters. This was based on group-average cluster analysis using a SIMPROF test on the available (albeit scant) data from all years (i.e. 1980s through to 2000s) that were predicted to be in an Estuarine/Marine state. This analysis was designed to indicate any extension to the current ecosystem state model that may be possible, by attempting to split cases in the terminal node for Estuarine/Marine further, rather than replace the current model (Figure 3.1).

Three terminal nodes were identified during CART analysis on the combined cluster solution (of four preliminary states, see Methods for additional detail, Figure 3.21). The relative cost of the model was 0.579, and the predictive success was 61% for both the learning and test data sets under cross-validation. The splitting variables for the model were the maximum daily tidal range and the minimum average water level.
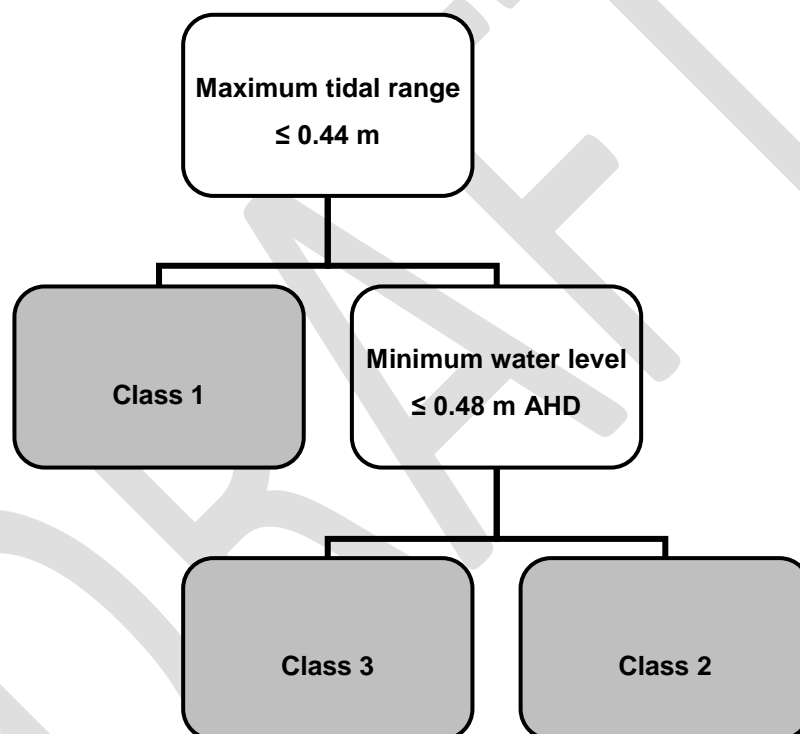
Figure 3.21. Classification tree based on four clusters identified amongst cases identified as Estuarine/Marine in the 1980s or the 1999-2007 data sets

ANOSIM analysis of the biological data for the cases within each terminal node was not significantly different for the data set including aquatic macrophytes and fish species presence/absence, nor for the data set including presence/absence data for aquatic macrophyte, commercial fish and invertebrate species. For the data set including bird and commercial fish species presence/absence, the overall ANOSIM result was non-significant (Global $R = 0.168$, $p = 0.086$), but pair-wise comparisons showed a significant difference between the first and second terminal nodes ($R = 0.302$, $p = 0.029$). Based on these analyses, there is no evidence that the CART model shown above adequately captures real differences identified in the biological communities of cases in the Estuarine/Marine state. This may be due to small sample sizes (given the p-value of 0.09 and the sample size of 18), to a lack of appropriate explanatory variables in the environmental data set available or be due to over-fitting of the biological data.

3.6. Sensitivity analyses

3.6.1. Sensitivity to initial clustering

Sensitivity to clustering technique

The clusters identified were somewhat sensitive to the method used to cluster the biological data set. For the long-term data set, three clusters were originally identified using the group-average clustering technique. When the membership of these clusters, was compared with those produced using *k*-means clustering of the same data ($k = 3$), the concordance was 85%.

The short-term data set including the fisheries-independent fish abundances was divided into four clusters using the original group-average clustering. The four *k*-means clusters identified had cases that were 84% concordant with those original clusters. For the second short-term data set (that including juvenile invertebrate abundances), there were originally five clusters identified. When the cases included in each cluster were compared with those calculated using the *k*-means clustering, the level of concordance was only 52%.

The effect of this sensitivity was then propagated through the CART analysis to determine the effect on the driving variables and thresholds identified. The *k*-means clusters did not produce the same original CART model as that obtained for the group-average clusters used in the original long-term model. The original long-term model had three terminal nodes, with average annual salinity and water depth from the previous year as the two splitting variables (Figure 3.22). The relative cost of the model was 0.384 and the predictive success under cross-validation of test samples was 81%. The model produced by the *k*-means clusters, however, had five terminal nodes at a relative cost of 0.181 (Figure 3.23). The first splitting variable remained the average annual salinity, but the second splitting variable was maximum turbidity, and subsequent splitting variables were the annual fluctuation in water level (annual maximum minus the annual minimum) and the annual maximum water depth. This model had a predictive success of 88% for the test sample under cross-validation. Neither maximum turbidity nor the annual fluctuation in water level was correlated with water depth from the previous year, but there was a strong correlation between the water depth from the previous year and the maximum depth ($r = 0.924$). Agreement between the cases assigned to each terminal node for the original group-average-based model and the *k*-means-based model was relatively high at 84%, with most of the discrepancy being due to the different number of terminal nodes identified by the two models.
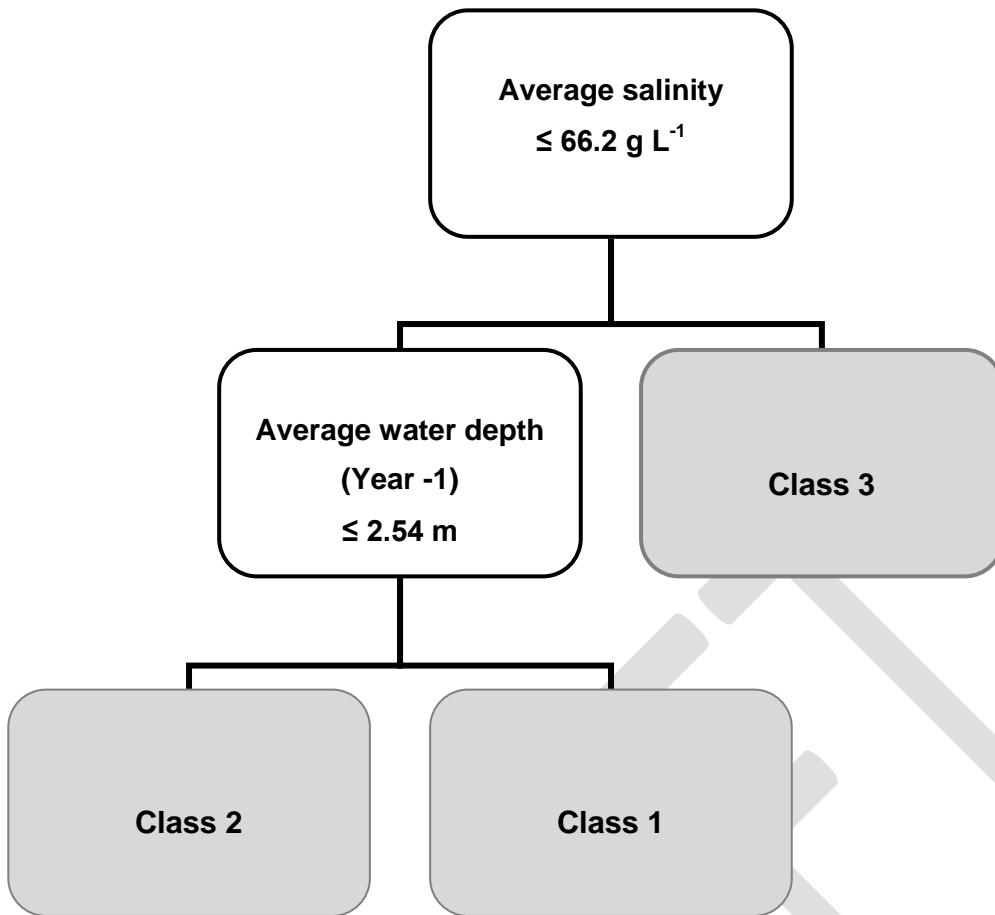
Figure 3.22. Classification tree based on group-average cluster analysis in the original long-term model development
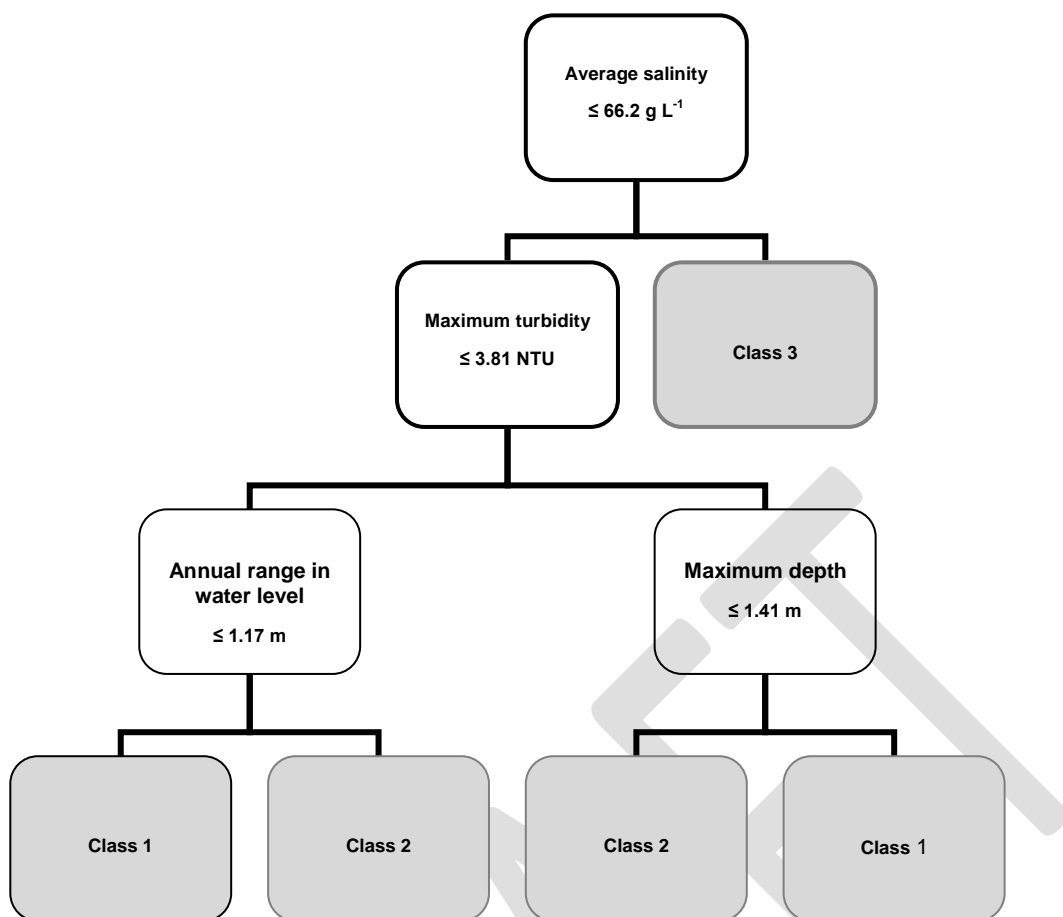
Figure 3.23. Classification tree based on k-means cluster analysis as an alternative to the original group-average clustering on the long-term data set

ANOSIM analysis tested the biological data associated with each terminal node arising from the *k*-means clustering. This indicated that there were only three distinct (i.e. statistically-significant) biotic groups described by the five terminal nodes. Re-assigning cases to these groups, and then re-analysing the environmental data using CART yielded a tree with four terminal nodes, determined by thresholds in the average annual salinity and the maximum turbidity and maximum water depth for each case. This model had a relative cost of 0.105 and a predictive success of 90% under cross-validation. Removing the maximum turbidity variable, along with other water quality parameters that were not able to be predicted, resulted in a tree with three terminal nodes at a relative cost of 0.271 (Figure 3.24). For this tree, the predictive success was 73% and the splitting variables were average annual salinity and maximum tidal range. Maximum tidal range was not significantly correlated with depth from the previous year ($r = 0.442$). Despite the differences in the resultant CART tree, the cases assigned to each state were 91% concordant between those produced using the *k*-means and group average clustering techniques.
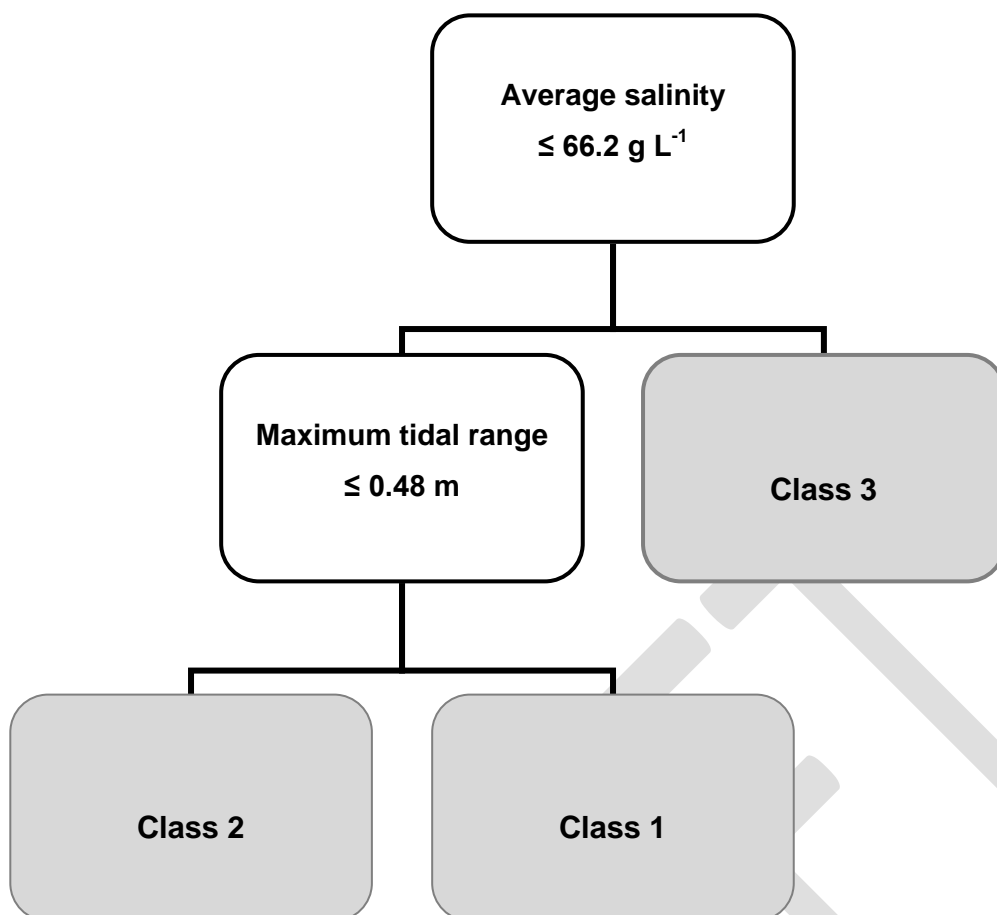
Figure 3.24. Classification tree based on k-means cluster analysis after combination of non-significantly different terminal nodes as an alternative to the original group-average clustering on the long-term data set

When the site-years in the Baseline scenario were assessed with the original group-average long-term model, 10% of site-years were predicted to be in State 1, 70% in State 2 and 20% in State 3. The model produced using the *k*-means clustering gave different proportions for States 1 and 2 (28% and 51%, respectively) although the proportion in State 3 remained 20%. Of the 1130 site-years in the scenario, the two models predicted concordant states for 80%.

Sensitivity to changes in the biological data set

The sensitivity of the cluster analysis to random changes within the data set was relatively low. Sensitivity increased with the number of clusters. For the long-term biological data, where three clusters were originally identified, changes of up to 50% of the original data set resulted in a high degree of concordance with the original clustering (mean > 82% for all proportions changed; Figure 3.25). The short-term biological data set including fisheries-independent fish abundances was grouped into four significant clusters originally (using group-average clustering). The relationship between the proportion of data changed at random and the percentage concordance with the original clustering was a negative linear relationship, with increasing standard deviation as the proportion altered increased (Figure 3.26). The short-term biological data set including juvenile invertebrate abundances originally yielded five distinct clusters. When this data set was altered and five new clusters were created, there was, in contrast, an exponential decay in the concordance of the new clusters with the original five (Figure 3.27). By the time 50% of the original data set had been randomly replaced, the agreement between the clusters was 54% on average.

These levels of concordance are all substantially higher than the theoretical minima for each number of clusters. For a three-cluster solution, a completely random generation of three numbers yields an average concordance of 0.40 over 1000 runs (SD = 0.03). A four-cluster

solution has a theoretical minimum of 38% concordance (based on 1000 runs of random numbers) with a standard deviation of 4%. Finally, for a five-cluster solution, all 1000 runs using random numbers yielded a 38% concordance with the solution based on the *k*-means clusters. All runs had higher levels of concordance than these obtained using only random number generation, indicating that some integrity remains in the data set despite the high proportion of data randomly changed for the final runs (up to 50%).
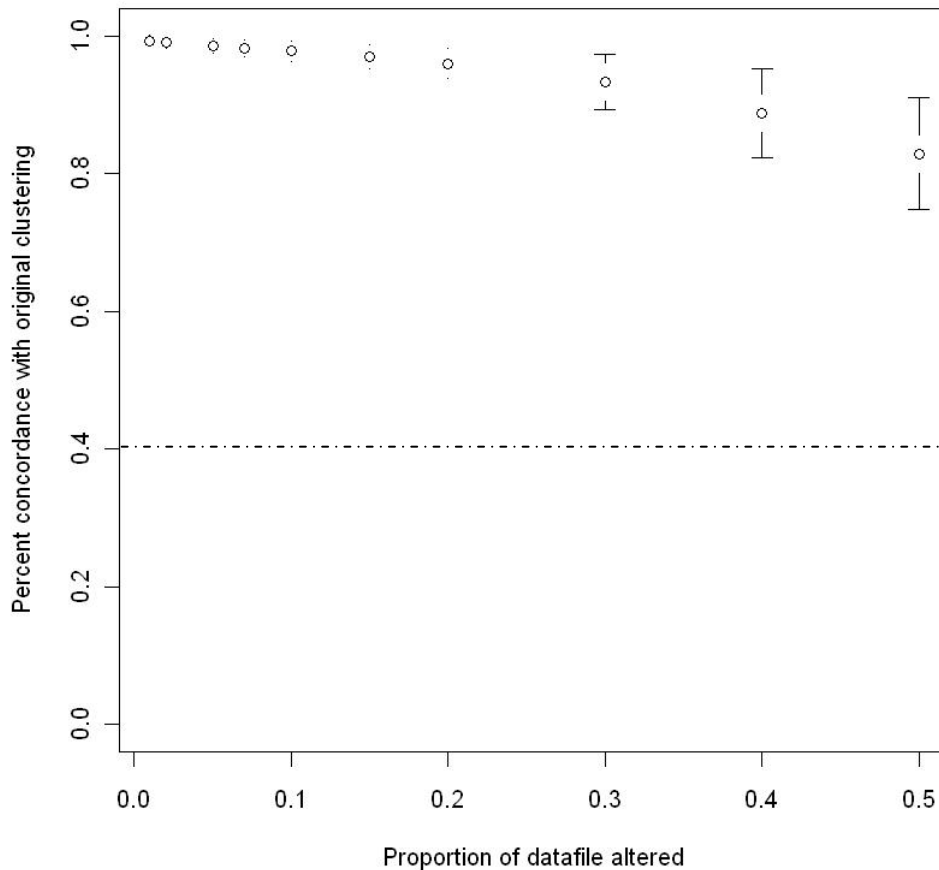
Figure 3.25. Sensitivity of cluster analysis to random changes within the long-term biological data set

Note: Points plotted are mean ± SE based on 1000 runs with increasing proportion of data set changed using resampling for negative values, calculated from a log(X+1) Bray-Curtis dissimilarity matrix. The dotted line indicates the theoretical minimum based on a completely random data set.
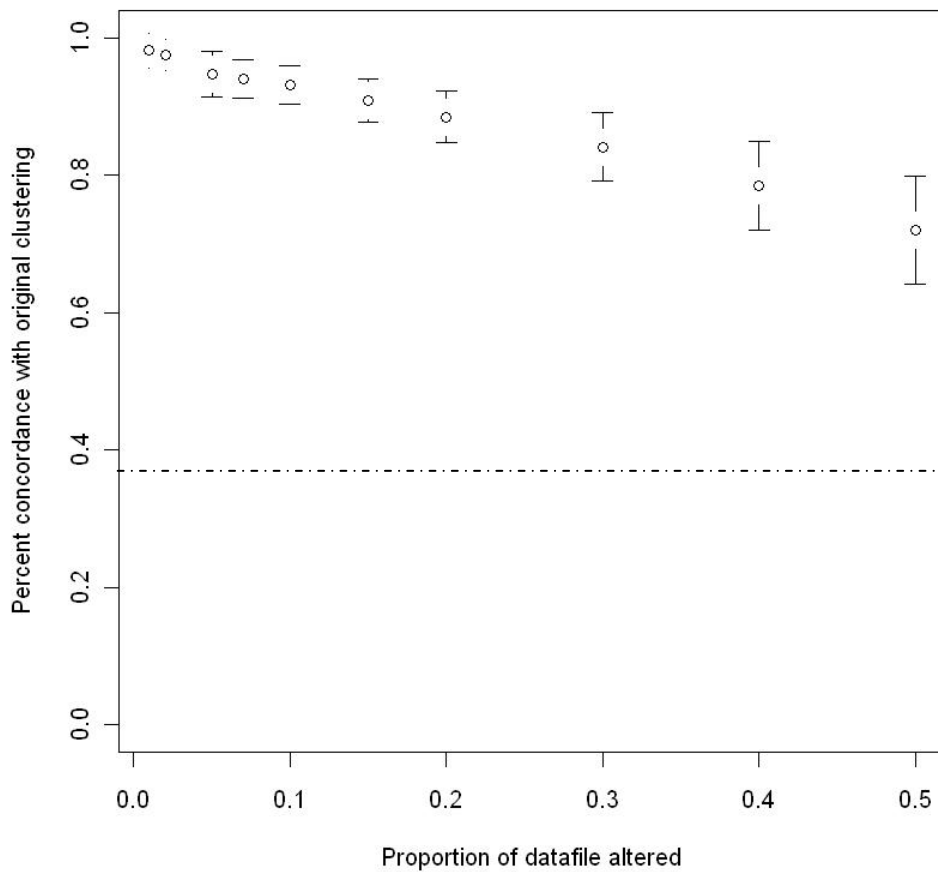
Figure 3.26. Sensitivity of cluster analysis to random changes within the short-term biological data set including fisheries-independent fish abundances

Note: Points plotted are mean ± SE based on 1000 runs with increasing proportion of data set changed using resampling for negative values, calculated from a log(X+1) Bray-Curtis dissimilarity matrix. The dotted line indicates the theoretical minimum based on a completely random data set.
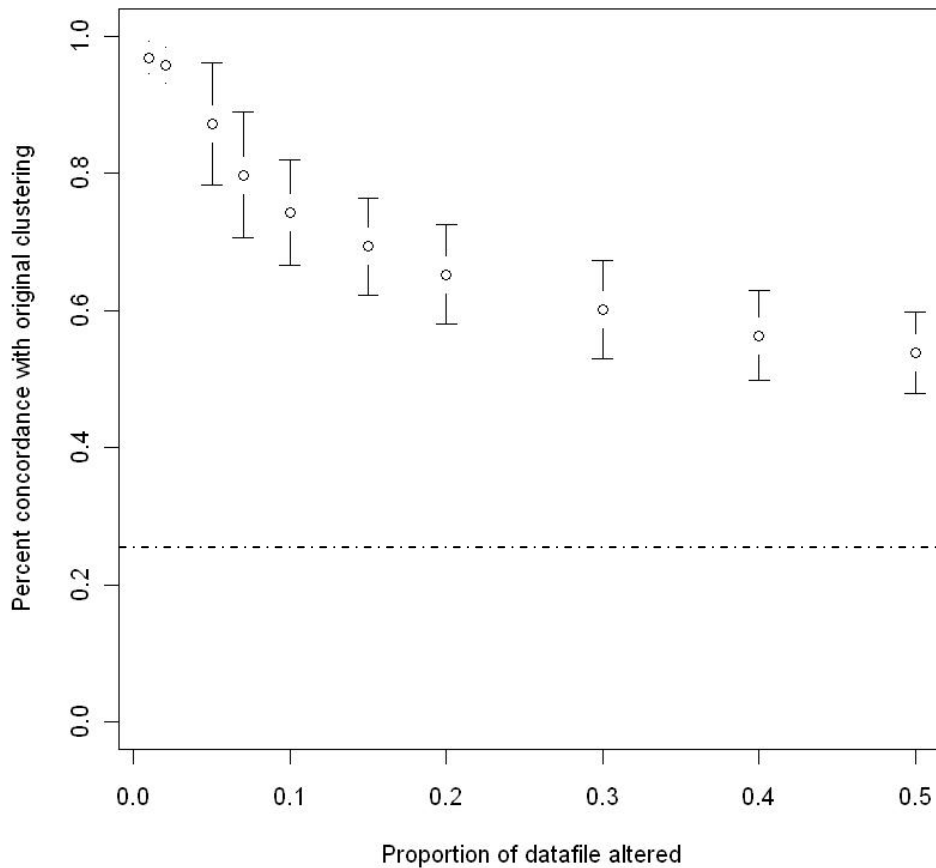
Figure 3.27. *Sensitivity of cluster analysis to random changes within the short-term biological data set including juvenile invertebrate abundances*

Note: Points plotted are mean ± SE based on 1000 runs with increasing proportion of data set changed using resampling for negative values, calculated from a log(X+1) Bray-Curtis dissimilarity matrix. The dotted line indicates the theoretical minimum based on a completely random data set.

When a random 20% of the biological data had been randomly changed prior to cluster analysis, there was relatively little change in the observed CART results. All five selected examples of 20% replacements had average annual salinity and maximum turbidity as their first two splitting variables. The overall number of terminal nodes varied between three and five (with two of five analyses yielding five terminal nodes, and the other three with three terminal nodes). The additional splitting variables were the annual fluctuation in water level and the maximum water depth (as was the case for the original *k*-means CART tree). Surrogate variables to the primary splitters (i.e. the secondary variables used for splitting when the actual splitting variables had missing values) were also very similar across all models.

For the five selected examples of a 50% substitution, the CART trees varied significantly from each other and from the original *k*-means tree produced. In all instances, only three terminal nodes were identified but the splitting variables included average salinity from two years' previous, average annual salinity, minimum water level, minimum depth, average depth from the previous year and maximum salinity.

The biotic assemblages for the terminal nodes of each model produced using substituted data were also analysed using ANOSIM. For each of the ten runs (i.e. five random examples of 20% data replacement and five of 50% data replacement), three distinct biotic groups were identified, regardless of whether the CART analysis had identified three or five terminal nodes. Once the relevant terminal nodes had been combined, the resultant states were 93% (SE = 2%) concordant with the original *k*-means states for the examples where 20% of the data had been substituted, and 83% (SE = 2%) concordant where 50% of the original data had been

substituted. One of the five runs where 20% had been substituted was re-analysed using CART, yielding a tree with three terminal nodes using average annual salinity and minimum water level as the splitting variables. The first of these splitters was identical to that identified by the original *k*-means analysis with identical thresholds for each, although minimum water level and depth from the previous year were not correlated significantly.

One example of substituted biological data (20% substitution) was used to predict states for the Baseline scenario site-years. This model predicted that 4% of site-years would be in State 1, 76% in State 2 and 20% in State 3. Both the *k*-means and the original group average-based models also predicted 20% of site-years in State 3, but the proportions for States 1 and 2 varied between models. The model based on substituted biological data had concordant predictions for 71% of site-years when compared to the *k*-means model, and 87% compared with the original group-average-based model.

Sensitivity to clustering environmental data instead of biological data

Group-average cluster analysis on the long-term environmental data revealed 16 significant clusters, while the equivalent clustering on the biological data resulted in only four significant clusters (one of which had a single case and was thus excluded as an outlier). If the group-averaged clusters for the environmental data set were truncated at the first four clusters, however, there was a reasonable degree of concordance in the cases assigned to each cluster (69%). *K*-means clustering (*k* = 3) on the environmental data set had a lower degree of concordance with the original group-averaged biological clusters, with only 58% of cases assigned to the same cluster. The k-means clusters on the environmental data had slightly higher concordance with the k-means clusters on the biological data, at 61%.

For the short-term data set, *k*-means clustering (*k* = 7) was used on the environmental data set and compared to the amalgamated clusters produced from group-average clustering of the two subsets of the short-term biological data (including juvenile invertebrate or fisheries-independent fish abundances). *K* was selected as 7 to correspond with the total number of clusters identified in the biological data set. Despite the large number of clusters identified, there was quite a high degree of concordance between the cases assigned by the two methods. A total of 79% of cases were assigned to the same clusters when the environmental data was analysed compared with the biological data.

3.6.2.   Sensitivity to threshold values

The first analysis of thresholds investigated whether the value of the threshold could be attributed to chance alone. When all thresholds were changed simultaneously, between 0.21 and 0.83 of site-years differed from the proportion expected in each state under a uniform distribution (Figure 3.28). Changing all thresholds together resulted in changes in distribution relative to a uniform distribution of states that was as extreme as or more extreme than the observed distribution in only 65 of the 1000 runs. This indicates that the observed distribution of states is not significantly different from the uniform distribution of states ($p = 0.065$), and that the distribution is relatively insensitive to the values of the thresholds.
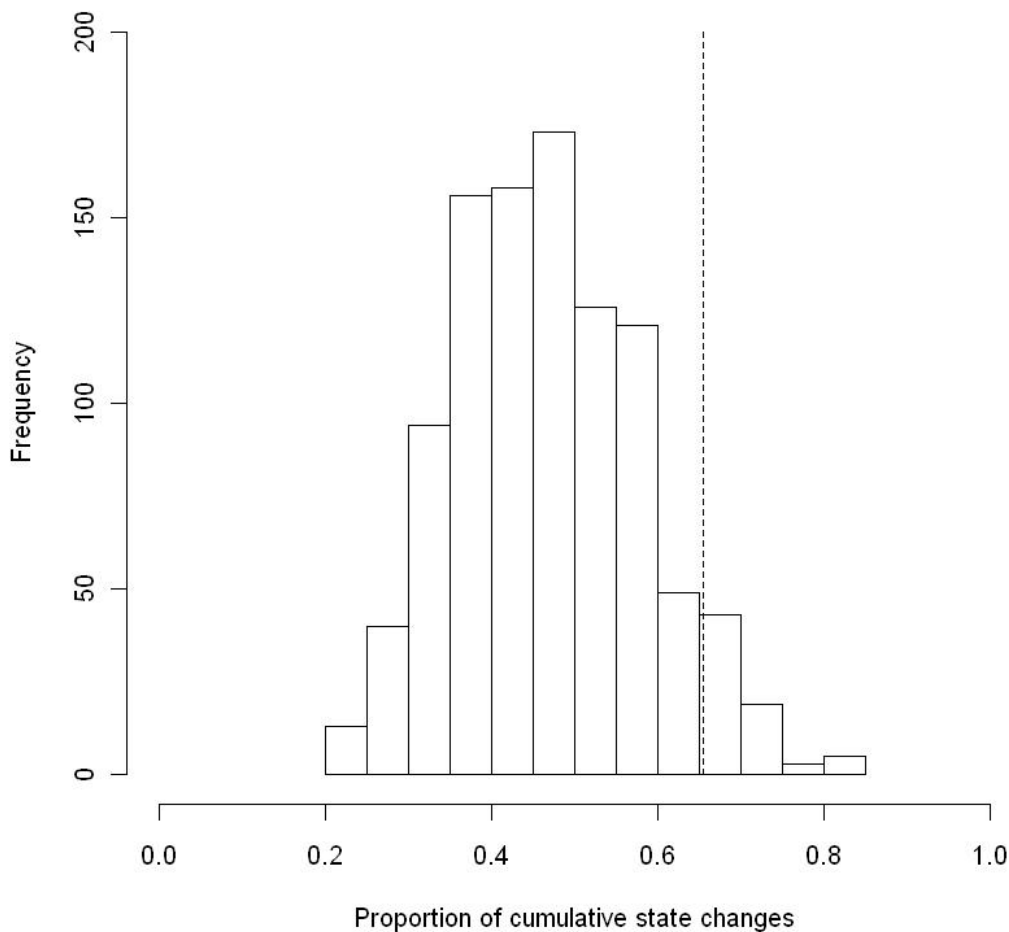
Figure 3.28. Histogram of change in state distribution relative to uniform with random changes to all thresholds

The dotted line indicates the observed value relative to a uniform distribution for the ecosystem state model.

The distribution of states predicted by the model was sensitive to the value of the threshold for tidal range (Figure 3.29). The change in state proportions relative to a uniform distribution for the 1000 runs varied between 0.56 and 0.81. Of these, however, 118 had distributions that were as extreme as or greater than that observed using the ecosystem model, meaning that the degree of sensitivity was not statistically significant ($p = 0.118$)
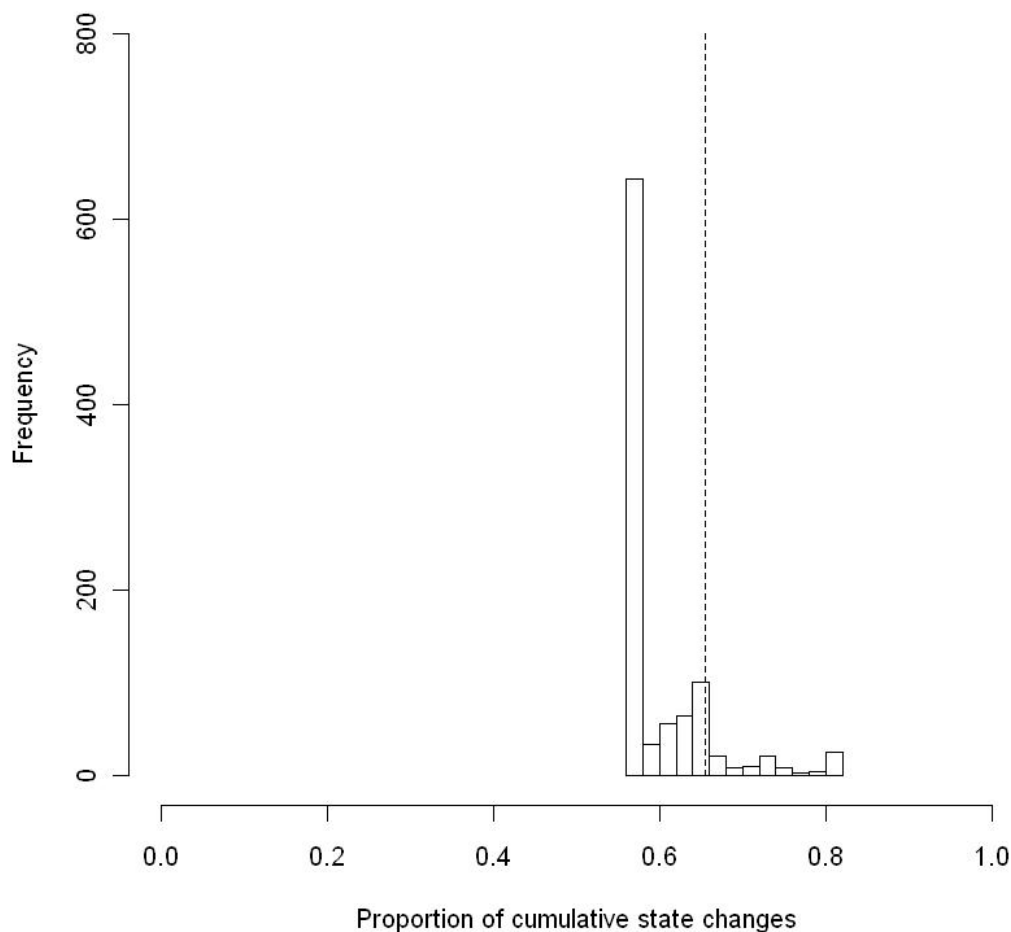
Figure 3.29. Histogram of change in state distribution relative to uniform with random changes to the tidal range threshold

The dotted line indicates the observed value relative to a uniform distribution for the ecosystem state model.

The same threshold for the maximum number of days without flow occurred twice within the model. We therefore tested changing the values for each separately and then both together. When the threshold for the maximum number of days without flow was varied for the marine basin only, the change in distribution of states was somewhat sensitive. The difference in proportions relative to a uniform distribution varied between 0.41 and 0.68 across the 1000 runs, with 122 runs showing a distribution as extreme as or more extreme than that observed using the ecosystem state model. Again, this was not a statistically significant result ($p =$ 0.122). Changing only the threshold for the hypersaline basin yielded similar results. The difference in state proportions ranged between 0.60 and 0.68, and 169 runs resulted in a more extreme or equally extreme distribution ($p = 0.169$). The sensitivity in the distribution of states was similar when the two thresholds were altered together (Figure 3.30). The proportion changing varied between 0.40 and 0.70, and 107 of the 1000 runs resulted in more extreme or equally extreme distributions. This was again a non-significant result ($p = 0.107$).
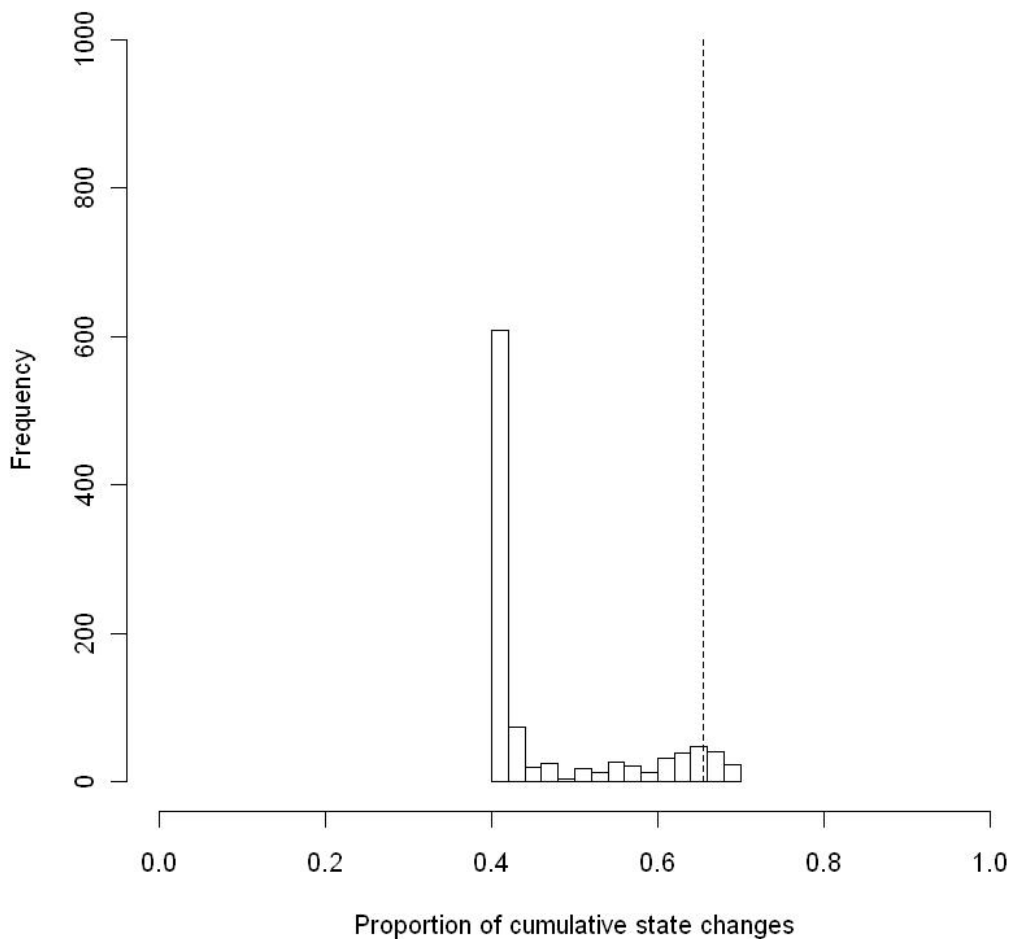
Figure 3.30. Histogram of change in state distribution relative to uniform with random changes to the maximum days without barrage flow threshold (changing both thresholds simultaneously)

The dotted line indicates the observed distribution relative to a uniform distribution for the ecosystem state model.

The first water level threshold (0.37 m AHD) had some impact on the distribution of states compared to a uniform distribution (Figure 3.31). The change in state proportions ranged from 0.60 to 0.69 across the 1000 runs. Of these 1000 runs, 304 had a distribution that was as extreme or more extreme than the observed distribution ($p = 0.304$), indicating that the observed value for the first water level threshold did not have a significant impact on the distribution of states predicted by the model.

Figure 3.31. Histogram of change in state distribution relative to uniform with random changes to the water level threshold at 0.37 m AHD

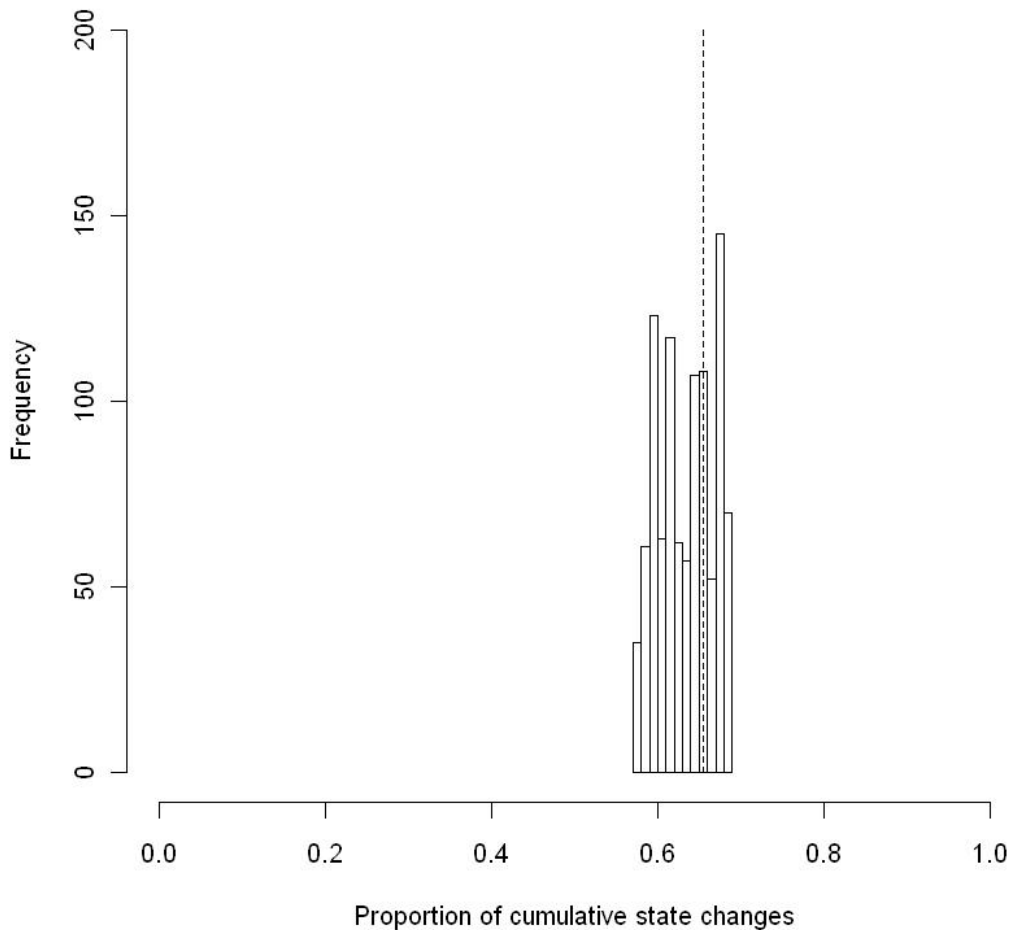The dotted line indicates the observed value relative to a uniform distribution for the ecosystem state model.

When the salinity threshold was replaced at random from the distribution of the salinity variable, there was no change in the proportion of site-years that differed from a uniform distribution across all states. The same was true for the depth threshold and for the second water-level threshold (-0.09 m AHD). This indicated that these thresholds were far enough down the decision tree to have no impact on the relative proportions of site-years attributed to each state compared with a uniform distribution.

The second analysis into the sensitivity of the threshold values investigated the effect of random changes on the threshold value on the predicted states relative to that predicted under the ecosystem state model.

When all thresholds were changed at random simultaneously, the average proportion of states that were different from the original model was 0.58 (Figure 3.32) (where 0 would be perfect concordance between the two sets of predictions and 1 would represent complete discordance). The coefficient of variation was 31%, indicating that there was a moderate degree of sensitivity in the number of site-years predicted to be in each state compared with the original model.
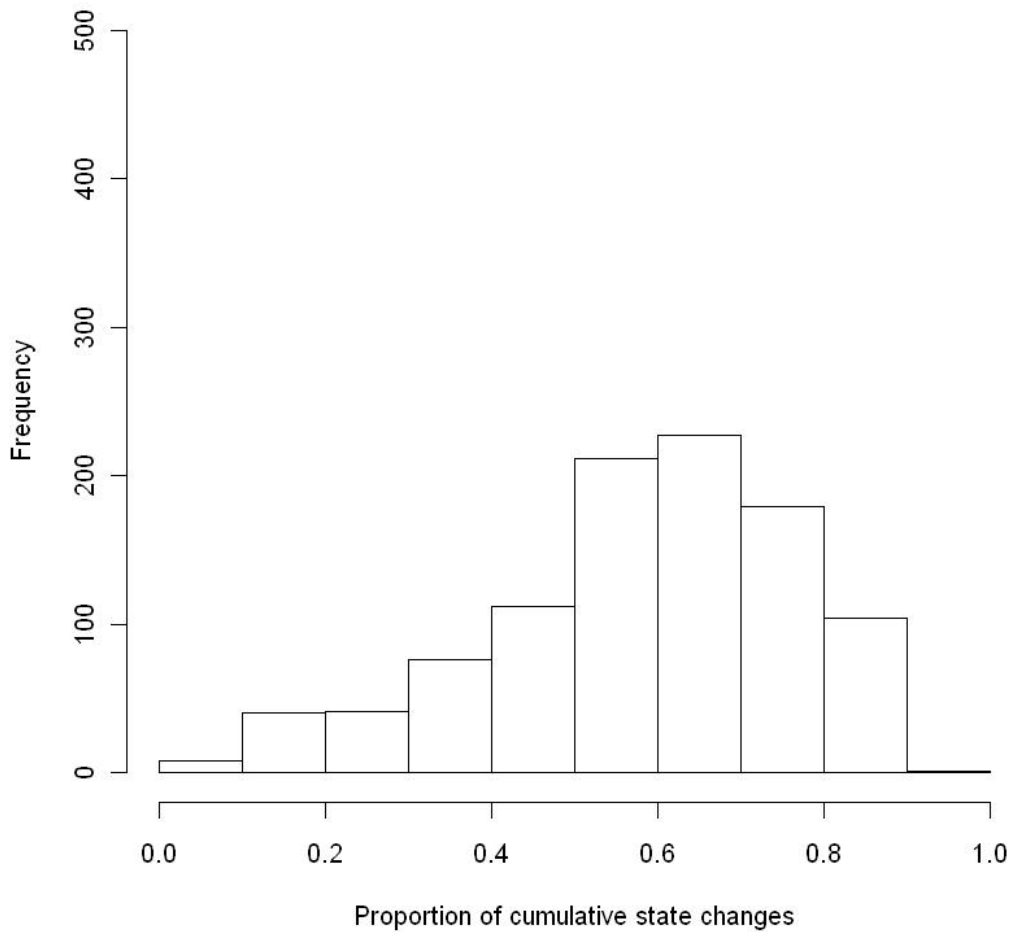
Figure 3.32. Histogram of change in state distribution relative to ecosystem state model predictions with random changes to all thresholds

The predictions for each site-year were sensitive to the value of the tidal threshold (Figure 3.33). Each run, on average 42% of site-years were allocated to a different state when the threshold was randomly changed (CV = 46%).
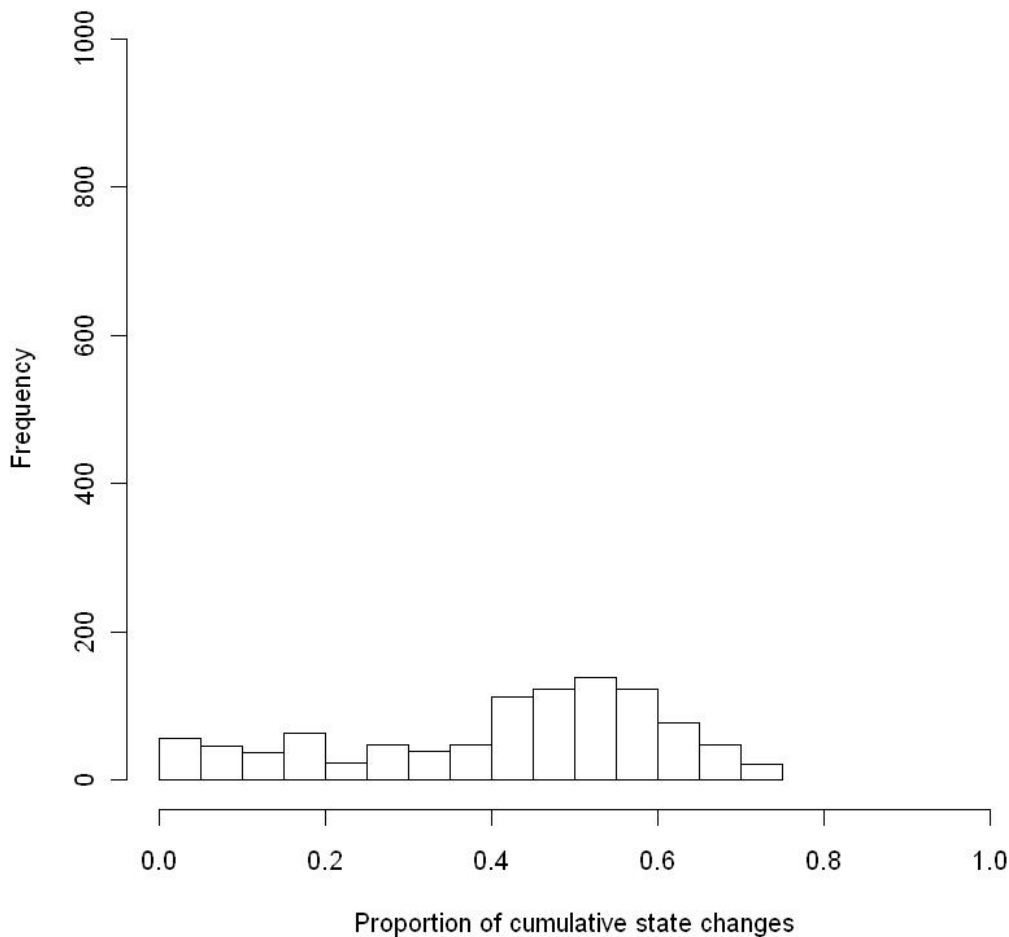
Figure 3.33. Histogram of change in state distribution relative to ecosystem state model predictions with random changes to the tidal range threshold

As before, the threshold for the maximum number of days without flow could be randomised for the marine basin, the hypersaline basin or both simultaneously. When the threshold in the marine basin was altered, the average proportion of site-years predicted to be in a different state from the ecosystem state model was 25% (CV = 71%). When the threshold in the hypersaline basin was changed independently, the resultant states were an average of only 10% different from those predicted under the ecosystem state model (CV = 60%). Both thresholds altered together produced state predictions that differed for 35% of site-years, on average, with a coefficient of variation of 69% (Figure 3.34).
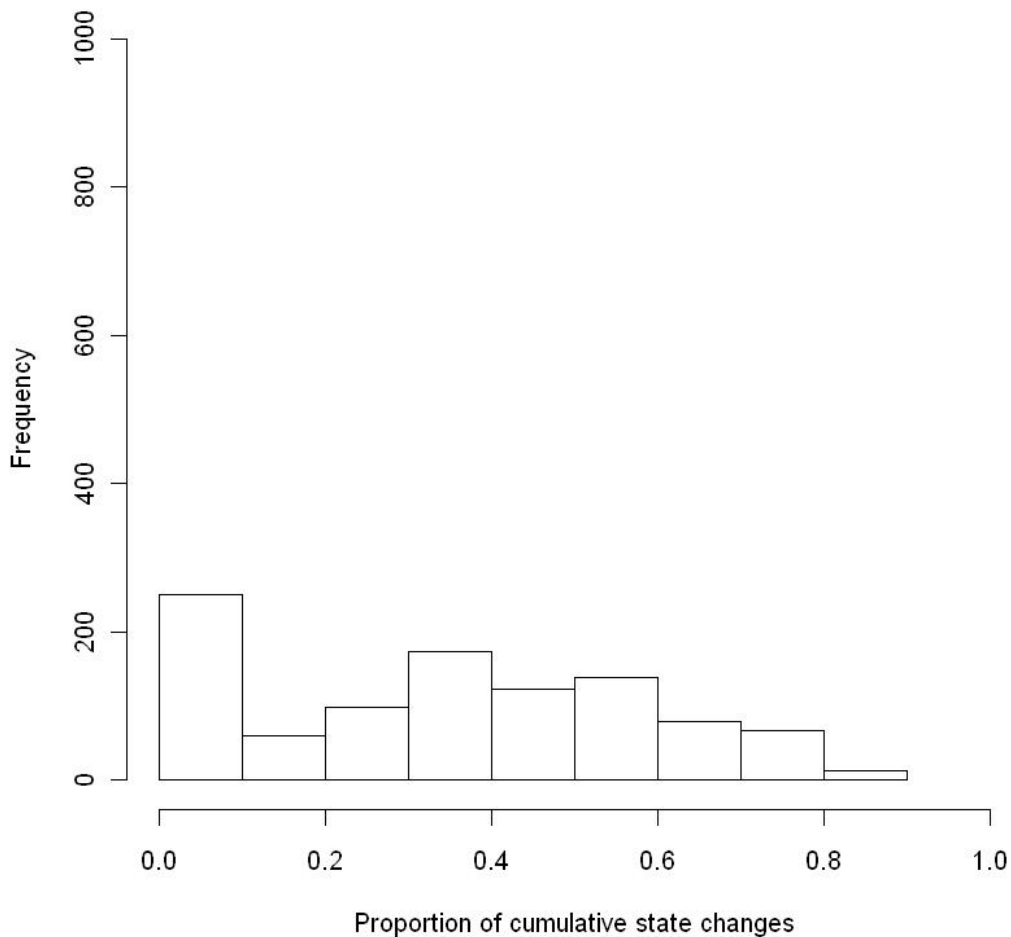
Figure 3.34. Histogram of change in state distribution relative to ecosystem state model predictions with random changes to the threshold for the maximum number of days without barrage flow

There was a small degree of sensitivity in the predicted states when the first water level threshold (0.37 m AHD) was randomly changed (Figure 3.35). On average, 6% of site-years were predicted to be in a different state from those identified under the ecosystem state model (CV = 88%). When the second water level threshold (-0.09 m AHD) was altered, however, there was very little change in the resulting predictions for each site-year (mean = 18%, CV = 22%)
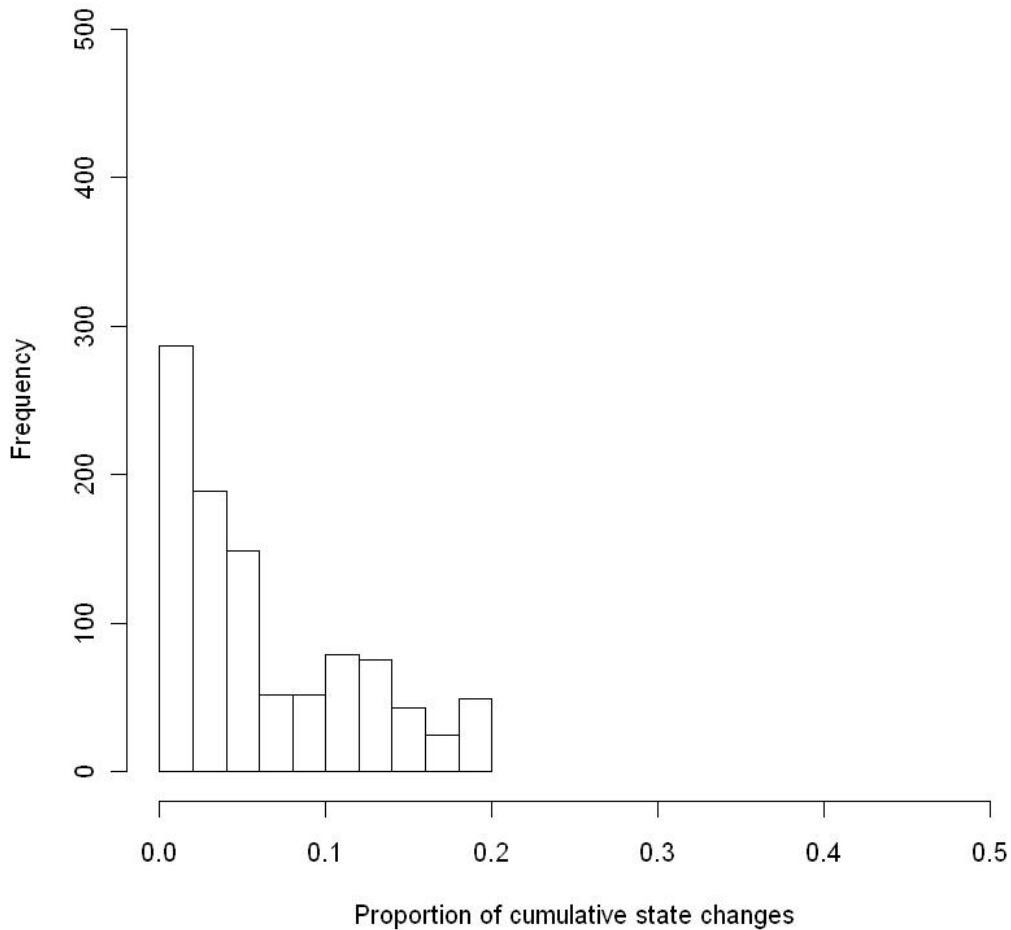
Figure 3.35. Histogram of change in state distribution relative to ecosystem state model predictions with random changes to the water level threshold at 0.37 m AHD

Note change in scale on x-axis from previous figures.

The proportion of states predicted to be in each state was quite robust to changes in the salinity threshold. Across the 1000 runs, the average change was 1% of site-years, with a coefficient of variation of 81%. The model was similarly robust to changes in the depth threshold, with less than 1% of site-years changing on average (CV = 107%).

4.   Discussion

4.1.   Model development

CLLAMM Futures was originally tasked with developing a single, ecosystem-scale response model for the Coorong that was capable of predicting the ecological response of the system to options for changes in management actions and climate change. We believe that we have been successful in fulfilling these requirements, overall. The model that we have developed is based on a truly multivariate analysis. Many other ecosystem response models in the past have been a collection of univariate responses (e.g. MFAT; Young *et al.*, 2003 or are based largely on literature reviews and untested expert opinions (e.g. many state-and-transition models; Petraitis and Dudgeon, 2004). The ecosystem state model described here is a data-driven, objective and multivariate approach to developing a state-and-transition model. It is not influenced by the 'pet' taxa of the scientists who constructed it, or by our preconceived notions of what is driving ecological change in the system (as far as is possible).

Using such an ecosystem state approach has a number of advantages. Firstly, there is significant interest amongst ecological scientists in alternative stable states (e.g. Petraitis and Dudgeon, 2005; van Nes and Scheffer, 2005 and in the identification of different ecosystem states (whether stable or not) and their ability to assist in the management of ecosystems (for example, Zweig and Kitchens, in press; Suding and Hobbs, 2009);, and this approach provides another tool to pursue these goals. Ecological changes in large, complex ecosystems, particularly given the uncertainty of climate change is also topical at the moment (for example in the west of the USA; Powell, 2008), and this approach provides a mechanism with which to predict these changes. Secondly, the concept has proved to be one that resonates with both the managers of the system and the general public. The notion that physico-chemical conditions are linked to the biotic communities that are found under those conditions, and that these change in space and time is intuitive and easy to explain to a non-technical audience. As a result, we have had broad acceptance of the concept and have been approached multiple times by a variety of groups to predict ecological responses to a variety of additional management scenarios. A major barrier to the adoption of many tools for managing ecosystems is the difficulty that can exist in conveying complex, and often highly mathematically-complex models to those stakeholders who the models are designed to assist (Zweig and Kitchens, in press). We have had no such issues in conveying the theory behind the ecosystem state model.

The ecosystem state model, as developed for the Coorong, is extremely well matched to the existing river and hydrodynamic models that exist for the region. The Murray-Darling Basin Authority has the capability to predict flows over the barrages under a wide range of past and possible future conditions. These data, along with climate change predictions that have been made by CSIRO (CSIRO, 2008), provide the input for the hydrodynamic model of the Coorong that is described in (Webster, 2006). All of the input variables required to use our model as a predictive tool can be constructed from the output of these two models. This largely occurred via happenstance. The model development process included many physico-chemical variables that could not be predicted by one of those two models (including water quality variables other than salinity and meteorological variables, for example), but these did not prove to be significant driving variables for the identified ecosystem states. This makes scenario analyses relatively simple to carry out for a wide range of possible future conditions. Another advantage of this occurrence is that, in the future, it will be relatively straightforward to monitor the potential ecological condition of the Coorong. Water levels and salinities are already logged at multiple locations in the Coorong, so managers will be able to assess their level of success in maintaining the ecological condition of the Coorong, should these ecosystem states be adopted as the basis for targets in the system. While we recommend ongoing biological monitoring to assist in the verification of model predictions, to identify causal links, and for the further refinement of the model, eventually this may not be necessary, and expensive biological

monitoring could be scaled back in favour of cheaper physico-chemical monitoring of the variables driving ecological responses.

We consider the alternating use of biological and environmental data sets in the model development process to be a strength of the approach. This allows for ongoing validation of the model as it is built, step by step. By doing this, we were continually assessing the links between the two types of data (which, from a theoretical ecological perspective, we would expect to exist in some form). It also provided a mechanism for detecting any trend towards over-fitting of any single component of the data, which is a tendency of CART modelling, in particular (Breiman *et al.*, 1984). This was apparent during the sensitivity analyses, where extraneous clusters or terminal nodes were detected and pruned, making the model quite robust.

Another strength of the approach is the use of multiple time steps in the final model construction. This means that both longer- and shorter-term changes are accounted for by the model, as was seen by the overall model's ability to capture the recent decline in the condition of the Coorong, which the initial annual model was unable to do. We also observed regular changes in state at both the seasonal and annual time-steps, indicating that the scale at which the model was built, and the scenario analyses were undertaken (i.e. annual), was appropriate.

The model development process was reasonably robust to missing and patchy data. The initial clustering step was the most sensitive to missing data. It was, however, possible to work around a certain amount of missing data. Some cases that were missing large amounts of data were excluded from the initial clustering step but could be used later to test the predictiveness of the model. We also used another approach during the short-term model development of conducting multiple cluster analyses and then aliasing the results across sub-sets of the data. Identifying environmental drivers of each biotic cluster provided an independent assessment of the success of this aliasing across analyses.

The environmental data set was much more robust to missing data. It was important however, during CART analysis to apply a penalty for the degree of missing data for each variable. This was because CART analyses tend to favour predictive variables with high proportions of missing values, because they produce fewer errors in classifying cases (Steinberg and Golovnya, 2007). In the end, the variables that were found to be driving ecosystem states tended to be ones that had sparse levels of missing data, but were also relatively easy to model.

While we were unable to include explicit spatiotemporal links in the predictive model to govern the allowable transition, such links are likely to occur in the Coorong between ecosystem states. Additional data is needed to test predictions to identify those transitions that do occur (but have not yet been observed) and to get an idea of the time required for any transition to take place (e.g. length of time for colonists to arrive). At the moment, all transitions are treated as possible and instantaneously occurring between time steps.

The ecosystem state model, as it stands, is heavily dependent on the number of days with no flow over the barrages in predicting which ecosystem states are present in the Coorong. For most scenarios, this is not a problem, because many managerial interventions and climate change are likely to affect the amount and duration of barrage flows. In times of low flow, however, drought management of the Coorong is often intended to foster ecological condition by mechanisms other than flow over the barrages. For example, the Murray Mouth is currently dredged to maintain a connection to the Southern Ocean (Murray-Darling Basin Commission, 2006). The original model was developed over a period of time where there were minimal engineering solutions occurring in the Coorong (i.e. Murray Mouth dredging was the only intervention, occurring from 2001 onwards), and so changes in ecological condition for the original model are closely linked to sources of fresh water. The alternative ecosystem model was developed to remove this reliance on River Murray flows and to give a more realistic assessment of the ecosystem states likely under these engineering-type solutions, particularly over the short term (see Lester *et al.,* 2009 for an application of this alternative model).

Caution should be used in interpreting the results from the alternative model, particularly relating to differences between the proportion of site-years predicted to be in either the Marine or Estuarine/Marine states, and the proportion of site-years predicted as Degraded Hypersaline,

because the relative proportions of both were different when the Baseline scenario was analysed using both models. Also, the predictions for individual site-years should not be emphasised, because there was quite a bit of variation in the prediction of the alternative model compared to the original model.

4.2.   Model evaluation

Model evaluation was attempted by comparing model predictions for the 1980s and 1990s with patchy data collected during those time periods and assessing the similarities between the biotic assemblages for each ecosystem state between the earlier periods and the time-span over which the model was constructed (1999-2007, which is referred to as the 2000s for convenience despite the presence of 1999). This was intended to assess how general the predictive model was, by comparing the predicted biota to the measured biota for a period outside the model domain. Some data were available for aquatic macrophyte presence/absence (including *Ruppia megacarpa*), coverage data for *Ruppia tuberosa*, bird abundances, catch per unit effort for the commercial fishery, and presence/absence of invertebrates. The years, the scale and locations over which these data were collected varied widely between data sets. Unfortunately there was very little overlap, making simultaneous comparisons involving multiple taxonomic groups (as was done during model development) very difficult due to limited sample sizes.

When each taxonomic group was considered separately, the various groups gave mixed results. Analyses for some groups detected differences in biotic assemblages between states, but others did not. Many analyses found significant differences between the biota supported by a given ecosystem state in the 1980s or the 1990s and the biota supported by the same state in the 2000s. The Estuarine/Marine state was the most common predicted for the Coorong in the 1980s and 1990s. Given the significant interaction between states and years across the various analyses, and the commonness of the Estuarine/Marine state compared with the other ecosystem states, the Estuarine/Marine state is the most likely to be significantly different between the 1980s and the 2000s. Thus this state may not represent a single ecosystem unit for the Coorong, and may be an amalgamation of several 'true' ecosystem states.

Multi-group comparisons were more consistent with the approach used to build the model in the first place. In comparing these groups, we expected that a good predictive ecosystem state model would produce significant differences between states, but not years, if the ecosystem states identified represented a complete set for the Coorong. These analyses also produced somewhat mixed results, largely due to the lack of overlap between the available data sets. Biotic assemblages in 1987 appeared to vary dramatically from that observed in the 2000s (based on an assessment of bird and commercial fish species present). Where there were data for more than one year in the 1980s, biotic assemblages tended to be grouped by ecosystem state, although, for the eight cases comparing aquatic macrophyte, invertebrate and fish assemblages, this was not statistically significant. This comparison did also indicate that there were differences between the 1980s and the 2000s, although that was not observed for a comparison of only aquatic macrophytes and fish.

Interpreting the trends from both the independent and combined taxonomic group comparisons, we believe that biotic assemblages do vary between ecosystem states, but that states present in the 1980s and 1990s may vary significantly from their counterparts in the 2000s. This is likely due to the environmental conditions in the earlier decades being so different from those experienced during the years over which the model was developed. As a result, we expect that the Estuarine/Marine state, in particular, is likely to vary widely in the biotic assemblages it supports.

One of the reasons for comparing model predictions to the 1980s, in particular, is that the Coorong was Ramsar-listed in 1985. This means that the ecological character of that time has been set as the de facto benchmark against which to measure current conditions within the wetland. The official ecological character description (Phillips and Muller, 2006) was not written until 2006, however, by which time significant changes in ecological character had occurred,

some of which are due to the current drought, but others are also linked to management during the intervening years. As such, there is not currently any reliable, detailed description of what the ecological character of the system was in 1985 at the time of Ramsar listing.

In order to assist in the assessment of scenarios, and to give context for management within the region, we have compiled a list of species known to be present in the Coorong in 1984 and/or 1985 (Appendix C). The most striking differences between the species present in the Coorong at the time of Ramsar listing and today are the loss of species from lower trophic levels. No bird or fish species that were recorded during 1984/85 have been lost from the system (although their distributions and abundances might be different now), but numerous macrophyte and invertebrate species have. This may reflect the lower mobility of plant and invertebrate species, and suggests that birds and fish reliant on these species for food and habitat have found suitable alternatives. This situation may not continue, however, with very few infaunal species remaining in the southern reaches of the Coorong (Rolston and Dittmann, 2009) and a sharp decline in the range and cover of the last remaining macrophyte, *Ruppia tuberosa* (Rogers and Paton, 2009).

A description of an example food web from the early 1980s is given by Geddes and Francis (2008) for Pelican Point. They indicate that the widespread presence of several macrophyte species (including *Ruppia megacarpa, Lepilaena sp.* and *Zostera muelleri*), along with beds of filamentous algae, would have been the basis for a large part of the food web. At the time, phytoplankton, microphytobenthos and macrophytes would have all contributed to the organic carbon in the system (Geddes and Francis, 2008). Today, however, the available organic carbon comes from largely planktonic sources (Geddes and Francis, 2008; Deegan *et al.*, 2009), particularly in the north of the system. The loss of this diversity in primary producers is likely to have contributed to the loss of macroinvertebrate species in the region. Species were lost from every macroinvertebrate class identified in the 1980s by the turn of the century, suggesting that the benchmark for ecological condition completed in 2006 (Phillips and Muller, 2006) was already representative of a degraded system relative to the early 1980s and that more work needs to be done to compile early data sets to gain a fuller appreciation for the biodiversity present at the time of Ramsar listing.

We also attempted to compare model predictions with data collected since the model has been developed. Again, very little data were available, with fisheries-independent fish-assemblage data the only data set we were able to obtain for 2008. Comparison of these data with data collected using the same method in 2006 and 2007 indicated that fish assemblages were significantly different again in 2008. Anecdotal evidence and the continued drought suggest that the ecological condition of the Coorong has continued to decline since the model was developed. Finding significant differences between the fish assemblages over such a short period of time may indicate that the ecosystem states of the Coorong have already moved outside the experience of the model. As conditions continue to decline, additional further degraded states may be needed to adequately predict the biological communities and conditions found there now. With the available data, however, such an extension of the ecosystem state model is not possible at this time.

Comparisons of biotic assemblages in the 2000s with those of the 1980s and 1990s did suggest that additional states in the Coorong are likely, particularly within the current Estuarine/Marine state. However, we did not successfully identify the environmental differences that drove these differences in biotic assemblage. CART analysis of the available environmental data for sites-years that were classified as Estuarine/Marine (across all years; 1981-2008) did not find driving variables and thresholds that resulted in biologically-distinct terminal nodes (which supports the ecosystem model as it stands). So while we may believe that the Estuarine/Marine state as identified within the ecosystem state model is a composite of several states, we have insufficient data in hand to resolve these differences and include them in scenario analyses, due to the small amount of additional data on which to base the analyses. As such, we have left the ecosystem model unchanged, with the caveat that the Estuarine/Marine state should be considered a conglomerate of states found in areas with a substantial tidal influence and with recent flows over the barrages.

4.3. Sensitivity analyses

Sensitivity analyses were conducted to identify areas of the original ecosystem state model that were sensitive to changes in method, or to values within the data set. These analyses allow an assessment of the robustness of the model, and highlight areas where future data gathering and model development efforts should be focused.

Analysis of the clustering technique used to identify the preliminary states (Step 2; Figure 2.2) showed that the more clusters that were identified, the lower the level of concordance between cases assigned to individual clusters using different clustering techniques. This is not an unexpected result. The more clusters that were present, the more possibility there was for change. Both the long-term (three clusters) and short-term (including fisheries-independent fish abundances; four clusters) data sets had quite high levels of concordance among clusters identified using group-average versus *k*-means clustering. The short-term data including juvenile invertebrate abundances (five clusters) was much lower.

The choice of clustering technique had little impact on the final ecosystem states developed by the model. This was due to the independent assessment of biological distinctness undertaken using ANOSIM. The CART models produced from the group-average and *k*-means clusters were quite different, but the models produced by the two methods were much more similar, once terminal nodes that did not have significantly-different biological communities associated with them had been identified using ANOSIM. Then, there was a high proportion of shared cases allocated to terminal nodes.

The degree of concordance between site-years assigned to each state dropped when the models were used to predict states for each site-year in the Baseline scenario. Here the concordance was only around 80%. A major source of error in the predictions for the site-years was our inability to predict turbidity (a variable that occurred consistently as a significant driver for *k*-means-based models). The *k*-means model with maximum turbidity as a predictor had a predictive success rate of 90% during cross-validation, but this dropped to 73% when maximum turbidity was excluded and replaced by maximum tidal range (a variable that we are able to predict).

It is impossible, therefore, to separate the loss of concordance under these Baseline scenario predictions that are due to the choice of clustering technique compared with that due to the drop in predictive accuracy due to the replacement of turbidity. The cross-validation results suggest that a predictive model including maximum turbidity would perform better than the one used and that the degree of concordance may be higher but, because we aren't able to predict turbidity, this remains unresolved.

The original rationale for using group-averaging clustering was to avoid making any *a priori* assessment of the number of ecosystem states that were likely to occur; however, if that number is high, then the clustering technique chosen will affect the cases assigned to each preliminary state. We are satisfied with this choice, despite the impact that this may have on the preliminary states that are identified. By happenstance, all of the predictive variables identified when group-average clusters were used were ones that could be predicted from the hydrodynamic model or inputs to the hydrodynamic model. However, the effect of the small differences in concordance amongst terminal nodes requires further investigation for future models to untangle the differences due to clustering technique from the differences arising from the replacement of turbidity in the model based on *k*-means clustering. This would be possible where models developed by the two clustering techniques involved only variables that were able to be predicted, so any differences would be due to the clustering technique, rather than variable substitution.

Sensitivity of the model to random errors within the biological data set was also affected by the number of clusters identified in the data set. As occurred for the choice of clustering technique, the greater the number of clusters, the more sensitive the model was to the proportion of error. All of the clusters investigated (three, four and five) were, however, significantly more robust to errors of even up to 50% than would be expected due to chance alone, indicating that the model overall is relatively insensitive to small numbers of random errors in the biological data set.

Another option for identifying preliminary states would have been to cluster the environmental data, rather than the biological data. Sensitivity analyses of the degree of concordance of doing this, compared with the clusters obtained from the biological data, led to mixed results. The long-term data set had a lower degree of concordance, despite a smaller number of clusters than the short-term data set. Concordance across seven clusters for the short-term data was unexpectedly high.

Clustering of the biological data set rather than the environmental data set was originally selected because we were primarily interested in the ecological entities rather than the physico-chemical entities that existed within the system. We would expect the biota to respond to the environment (rather than the other way around, most of the time), so we treated the biotic assemblages as the dependent variable for the CART analyses. Also, management agencies tend to measure salinity, and other physico-chemical parameters, when assessing the environmental condition of systems like the Coorong, so it made sense to investigate them as drivers or indicators of ecosystem change. The ecosystem state model as it stands, is also one that can be routinely modelled by MDBA (with the inclusion of the Coorong hydrodynamic model in BigMod), so approaching the modelling in this direction meant that the end result is one that is more likely to be of use to managers and other stakeholders.

The distribution of states predicted by the ecosystem state model was most sensitive to threshold values for splits at the top of the decision tree. This means that the model was more sensitive to the exact values of threshold for tides and maximum days without flow, but less sensitive to the exact value of the water level, depth or salinity thresholds. Thresholds for variables near the bottom of the decision tree had very little impact on the overall distribution of states predicted by the model. Overall, however, the distribution of states was not significantly different from a uniform distribution despite changes to all thresholds and each threshold individually.

It was quite surprising that the model predictions were so insensitive to changes in the salinity threshold. Salinity was the variable considered by most experts to be driving the ecological changes in the Coorong in recent years. Our findings, however, suggest that this is only a small part of the overall story, with the lack of freshwater flows and water levels playing a much larger role in driving the identified ecosystem states. This may be partly due to the close links between freshwater flows and changes in salinity. Freshwater flows will freshen the Coorong, but will also bring inputs of particulate organic carbon and nutrients, amongst other things, which may be important for biogeochemistry in the Coorong (Cook *et al.*, 2008), and barrage flows also play an important role in increasing water levels throughout the Coorong (Webster, 2005). This combination of benefits may be more important for driving the ecological condition of the Coorong than a simple dilution of salinity within the system.

Another analysis that could be run would be to compare the existing model with one developed from a data set excluding rare species. The opposite approach could also be taken, with a model developed from a data set of just the presence/absence of species, rather than abundance data. Both of these analyses would test the generality and robustness of the model constructed and provide clues to the applicability of the model, as constructed, to other similar systems. This may also point the way to how the routine collection of ecological data within the Coorong could be optimised, i.e. made more cost-effective and efficient, through streamlining, and hence more affordable.

4.4. Limitations of the modelling

The major limitation of the Coorong ecosystem states model is that it is biased towards describing declines in ecosystem condition, rather than any recovery that may occur. The data set upon which the model was based (from 1999 to 2007) included only relatively dry years and low flow conditions. Small barrage releases occurred towards the beginning of the time series, but for the last few years, only fishway flows, or no flows at all, were observed through the barrages. This is a highly unusual occurrence for the Coorong, as is demonstrated by the scenario analyses undertaken (Lester *et al.*, 2009). As such, the ecosystem states that have

been identified, as well as the transitions that govern changes between them, are focused upon describing poor environmental conditions and ongoing decline of the ecosystem. This is apparent when the fact that five of the eight identified ecosystem states appear in the Coorong only after 11 continuous months with no barrage flows – a situation that was largely unknown in the Coorong until recent years. As a result, there is significant uncertainty about the ability of the model to predict any recovery within the system, or to categorise ecosystem states for more estuarine conditions. This is highlighted by the analyses comparing the 1980s to the training data set, where the biota of states in the 1980s was frequently significantly different from the biota of the same state in the 2000s.

It is likely that there will be hysteresis within the Coorong ecosystems, meaning that the pathways for recovery will differ from the pathways for decline. There is also likely to be time lags inherent in the recovery of the ecosystem after flow return. Recovery from drought is known to take significantly longer than recovery from flooding in freshwater systems (Lake, 2000), and the same is likely to occur for estuaries. These unknowns associated with recovery mean that we are unsure as to the ability of the ecosystem model to predict recovery accurately, both in terms of the timeframes required and the transitional stages along the way. This will only be resolved if data are collected during recovery phases and the predictions of the model are able to be tested and refined accordingly.

Despite this limitation, the ecosystem state model is still useful for predicting ecosystem responses to changes in management and climatic conditions. The ecosystem states described can be thought of as being the worst case, and the 'Healthy' and 'Average' states identified should be considered the de facto baseline for ecosystem condition, with additional ecosystem states possible and likely beyond these. The current condition of the Coorong is such that even reaching the healthier end of the spectrum of described states will take time and effort, so there should be adequate time to refine the model before those predictions are likely to be needed.

It should be stressed that the model produced does not necessarily rely on having captured causal relationships. The CART algorithm seeks the best variables and values thereof to split the cases into a series of end members that represent the identified states. Thus the threshold values of the chosen variables are those that are associated with the various end members. It is possible to arrive at more than one pathway to a given fate, as was seen in the alternative model for the South Lagoon.

The use then of those splits in our state-and-transition model is our attempt to use environmental conditions to represent what is happening to the biota in the Coorong by association with key physico-chemical measures. This was deliberately chosen because the managers of the Coorong (as for many waterbodies elsewhere) are able to routinely measure the waters for physico-chemical variables much more easily than for biological or ecological variables. Our use of those variables chosen by the CART algorithm to indicate different ecosystem states was tested in the sense that the states had to represent different collections of biota and were shown to be robust to analytical decisions along the way by the re-testing steps, validation and evaluation done as part of the model development. As such we suggest that the key splitting variables and their threshold values be utilised as indicators of the ecosystem states that they are associated with in the model.

It is a separate but interesting question as to whether the model does capture information about the states that could be interpreted causally. That would require further testing that should be possible in the future under different environmental conditions and with judicious use of manipulative experiments. The easiest way of testing the ideas encapsulated in the model is to predict the ecosystem states to be seen within the Coorong under a wider range of future conditions – essentially what has been done under the scenario modelling (Lester *et al.*, 2009). But the test of those predictions will be to monitor not only the states of the future (i.e. the biotic elements of each site) but also the five or so key environmental variables shown in the model splits. The test may be imperfect because of unknown (as yet) trajectories of recovery or further transitions to state not yet in the overall model (as discussed above).

More direct testing will involve experiments but these may need to be set up the field rather than the laboratory. Obviously in times of drought we cannot experiment with the release of water

across the barrages. We had hoped during the formative stages of the CLLAMMecology Research Cluster in build some testing into the research program but alas there were no barrage releases during these years. Instead experimental protocols were developed for when such releases might become possible sometime into the future. We still need to derive explicit predictions as hypotheses from the model and the other themes within CLLAMMecology to make this a complete experimental design for future experiment and then the fieldwork teams need to be marshalled at that time and for some return times after the release (to check for delayed responses). Apart from those large-scale management experiments, it may also be possible to set up mesocosms within the Coorong lagoons to test the effects of particular water depths or quality for the development of the different ecosystems and their biota. Those smaller-scale experiments may be particularly worthwhile for determining how quickly transitions between states might occur and the relative ease of particular transitions that lead to recovery. This would need to be done in the field to maximise the chance of recruitment of biota to the experimental units but we would have to monitor the availability of propagules and other chance occurrences during the experiment to interpret the outcomes properly (Petraitis and Dudgeon, 2004; Petraitis and Dudgeon, 2005).

4.5.   Management implications

This approach of defining ecosystem states has significant implications for the research and management of estuaries and other large, complex ecosystems. The multivariate data input and exploratory nature of the technique has the advantage of greatly simplifying the task of defining ecosystem condition and has the potential to revolutionise the setting of management targets. This approach allows management of a system at an ecosystem scale. Management could now use a mix of appropriate ecosystem states as a management objective, rather than using a few surrogate variables or indicator species that are often arbitrarily chosen or defined and may lack adequate testing. The ecosystem states described here are combinations of environmental conditions and associated biota as determined by the data in hand, rather than any preconceptions of co-occurring focal variables. By managing for a set of states, there is significantly more flexibility in the definitions of limits of acceptable change. Another major advantage is that functional redundancy in an ecosystem, that is where multiple species could perform the same function, is recognised and incorporated into the ecosystem state model, so is accounted for in the management strategy, rather than being overlooked, or even penalised as inappropriate change. Using a multivariate approach highlights the interconnectedness and complexity of the ecosystem as a whole and therefore discourages the temptation to search for a 'magic bullet' solution to ecological problems (Lester *et al.*, in prep). This means that management focus is more likely to remain on finding feasible solutions to ecological problems, rather than chasing unlikely quick-fixes.

An important role for management arises from the need to further validate the ecosystem state model predictions. Tailored monitoring plans focused on the biota of the Coorong (particularly the fish, birds, macrophytes and invertebrates upon which this model is based), along with the hydrodynamic variables that the model suggests drive changes in biotic assemblages will provide additional data for the ongoing testing and refinement of these models. This will be particularly important through any interventions that are carried out, with monitoring needed before, during and for several years after any actions designed to affect the ecological condition of the Coorong.

In the interests of parsimony, we investigated whether any biotic groups could be excluded from biological modelling. However, even looking at only the top five species distinguishing amongst ecosystem states (keeping in mind that multiple sub-sets of data were used), all taxonomic groups and a wide variety of species within each were important for distinguishing between ecosystem states. The presence of macrophytes, bird abundances (including waders, waterfowl and fishing birds), both commercially-caught and small-bodied fish abundances (or CPUE) and invertebrate abundances (including polychaetes, amphipods and chironomids) were all useful indicators of the ecosystem state of the Coorong system. This suggests that full biological surveys should continue, at least until model predictions have been validated further

and any necessary model refinement has occurred. It does, however, also justify the multivariate nature of the ecosystem state model, indicating that constructing the model based on a single taxonomic group would not be as effective at identifying ecosystem states as including all available data, as was done here.

In addition to monitoring the variables identified here as driving the ecosystem states of the Coorong, we also recommend that water quality variables continue to be measured, as often as possible. One of the reasons for the lack of water quality variables identified as driving variables may be the resolution of the existing water quality data. The best data on water quality (other than salinity) are collected by DEH as spot samples from a handful of locations at quarterly intervals. Logged data on a daily time step, for example, may allow us to identify any additional relationships between these variables and any changes in biotic assemblages. Turbidity, for example, regularly appeared as a driving variable in the sensitivity analyses, and better data may a) allow us to resolve the role that turbidity does play, and b) develop the ability to predict turbidity in the Coorong in the future. In addition, it is likely that ecological conditions have continued to decline since the model was developed (as is suggested by the comparison between fish abundances in 2008 to those in the training data set), and water quality variables may play an important role in defining those changes that were not apparent in the model to date.

4.6.   Applicability of ecosystem state modelling to other systems

The development of ecosystem state models, such as this, should be possible for a wide range of aquatic systems. Many systems have a tendency to experience stepped changes in condition (e.g. lakes, wetlands, estuaries; Folke *et al.*, 2004). For example, step changes occur where wetting and drying cycles are present, or where ecosystem-engineering species exist, whose presence or absence dramatically changes the ecological character of a system (Folke *et al.*, 2004). The basis for modelling our ecosystem states is the link between biotic assemblages and thresholds in the surrounding physico-chemical conditions. This link occurs in the vast majority of aquatic systems, so the approach should be applicable in each. Each new system, however, will need to be explicitly and empirically modelled until an understanding is gained about the generality of the relationships observed (and some assessment of causality). At this stage in its development, we believe it is certainly worth applying ecosystem state modelling elsewhere, if only to demonstrate its generality, or otherwise. Should such modelling be as successful elsewhere as in the Coorong, it may prove to be a significant step forward in our understanding of, and ability to manage, environmental assets at an ecosystem scale.

5. Conclusions

We were successful in building a single model that can describe the probable ecological responses of the Coorong to a range of 'levers' including both management actions and likely consequences of climate change. The model, as it stands, is intuitive and useful to managers, with its identified combination of water regime and salinity drivers.

In building the ecosystem state model, there have been many innovations and extensions of previous ecological modelling approaches. We modified the theory of alternative stable states to make it more practical and relevant to dynamic and degrading ecosystems. We used a chain-of-methods approach on a broad, disparate aggregate data set that included multiple taxonomic groups and a plethora of possible physico-chemical variables (not simply limited to their means). The tests that have been used to test the individual steps in the model development process, the sensitivity of the model and the validity of its predictions have not, to our knowledge, been applied in this manner before. We have also spent an extensive amount of time and effort in the development of novel ways to display the outputs of the modelling and the scenario analyses in a manner that is comprehensible, we hope, for both scientific and lay audiences.

We expect that this approach will lead to ongoing improvements in modelling of ecological responses in complex ecosystems, and in the management of these systems. We believe that there are many exciting future directions arising from this work, including addressing some of the identified limitations of the current model, but also in applying this modelling approach to other ecosystems (both nearby and distant), and working with managers to apply the results to an adaptive management framework for the Coorong.
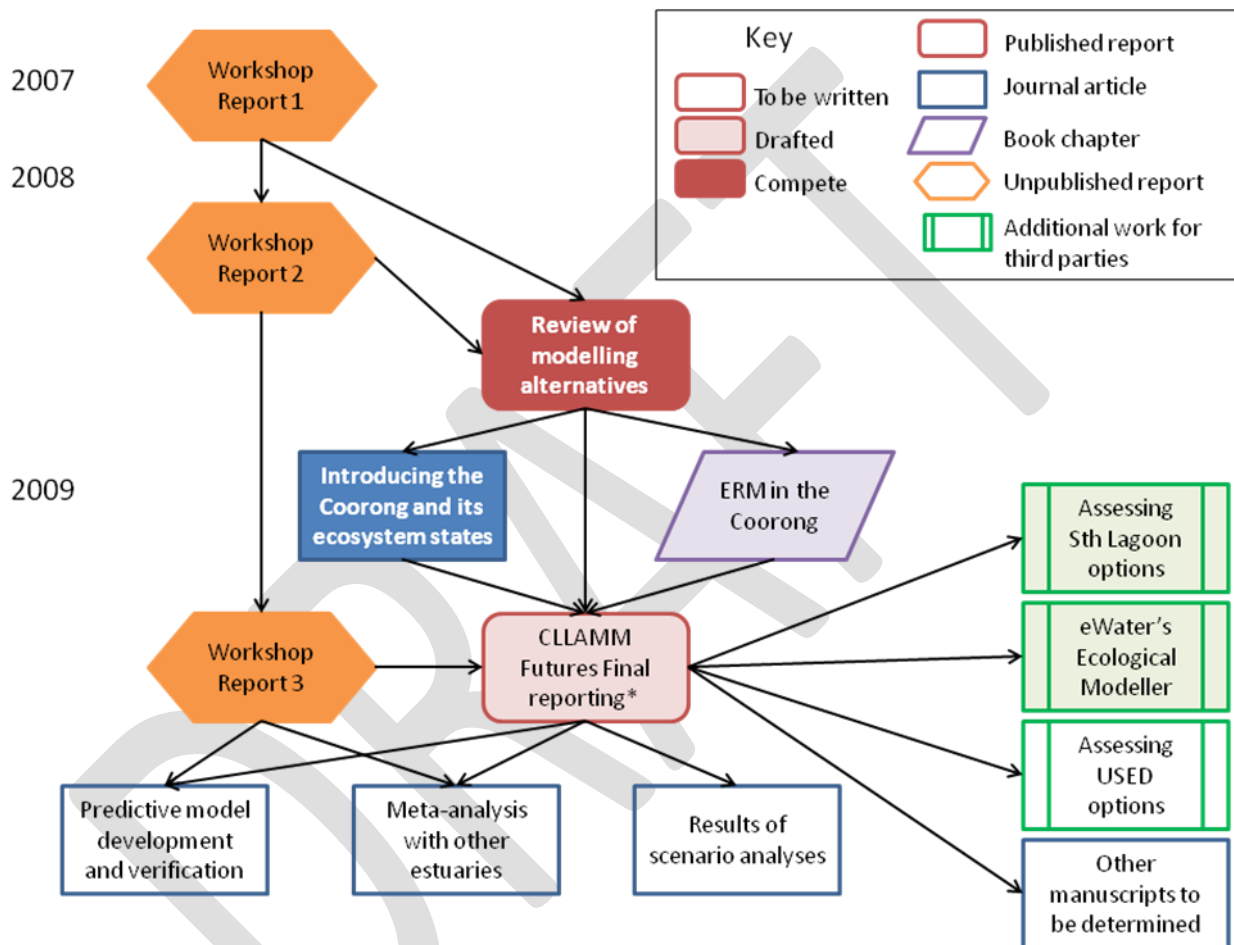
6.   References

Anderson, M. J., Gorley, R. N. and Clarke, K. R. (2008) PERMANOVA+ for PRIMER: Guide to software and statistical methods, PRIMER-E Ltd, Plymouth.

Bestelmeyer, B. T., Herrick, J. E., Brown, J. R., Trujillo, D. A. and Havstad, K. M. (2004) Land management in the American Southwest: A state-and-transition approach to ecosystem complexity. Environmental Management, 34, 38-51.

Breiman, L., Freidman, J. H., Olshen, R. A. and Stone, C. J. (1984) Classification and regression trees, Wadsworth, Belmont.

Clarke, K. R. (1993) Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology, 18, 117-143.

Clarke, K. R. and Gorley, R. N. (2006) PRIMER v6: User Manual/Tutorial, PRIMER-E Ltd, Plymouth.

Close, A. and Sharma, P. (2005) Hydrologic models used for water resource and salinity management in the River Murray system. In 28th International Hydrology and Water Resources Symposium (Ed, Institution of Engineers) Wollongong, NSW.

Cook, P., Aldridge, K., Lamontagne, S. and Brookes, J. (2008) Element and nutrient mass-balances in a large semi-arid riverine lake system (the Lower Lakes, South Australia). CSIRO Water for a Healthy Country National Research Flagship, Canberra.

CSIRO (2008) Water availability in the Murray: A report to the Australian Government from CSIRO Murray-Darling Basin Sustainable Yield Project, CSIRO, Canberra.

De'ath, G. (2002) Multivariate regression trees: A new technique for modeling species-environment relationships. Ecology, 83, 1105-1117.

De'ath, G. and Fabricius, K. E. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, 81, 3178-3192.

Deegan, B. M., Lamontagne, S., Aldridge, K. T. and Brookes, J.D. (2009) Trophodynamics of the Coorong. Food Webs: Their baselines, structures and dynamics in a disturbed aquatic ecosystem, CSIRO Water for a Healthy Country National Research Flagship, Canberra.

Department for Environment and Heritage (2000) Coorong and Lakes Alexandrina and Albert Ramsar Management Plan, South Australian Department for Environment and Heritage, Adelaide.

Dittmann, S., Cantin, A., Imgraben, S. and Ramsdale, T. (2006) Macrobenthic Survey 2005: Murray Mouth, Coorong and Lower Lakes Ramsar site. Report for the Department for Environment and Heritage, Adelaide.

Dittmann, S., Dutton, A. and Earl, J. (2008) Macrobenthic survey 2007: Murray Mouth, Coorong and Lower Lakes Ramsar Site, South Australian Department for Environment and Heritage, Adelaide.

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L. and Holling, C. S. (2004) Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology, Evolution and Systematics, 35, 557-581.

Gaines, S. D. and Denny, M. W. (1993) The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. Ecology, 74, 1677-1692.

Geddes, M. (1987) Changes in salinity and in the distribution of macrophytes, macrobenthos and fish in the Coorong lagoons, South Australia, following a period of river murray flow. Transactions of the Royal Society of South Australia, 111, 173-181.

Geddes, M. and Butler, A. (1984) Physiochemical and biological studies on the Coorong Lagoons, South Australia, and the effect of salinity on the distribution of the macrobenthos. Transactions of the Royal Society of South Australia, 108, 51-62.

Geddes, M. C. and Francis, J. (2008) Trophic ecology pilot study in the River Murray estuary at Pelican Point. SARDI Aquatic Sciences Publication No F2007/001193-1. Report Series No 274.

Gosbell, K. and Christie, M. (2004) Wader surveys in the Coorong and SE Coastal Lakes. Australasian Wader Studies Group.

Horn, H. S. (1975) Markovian properties of forest succession. In Ecology and Evolution of Communities(Eds, M., Cody and J., Diamond) Harvard University Press, Cambridge, MA, pp. 196–211.

Lake, P. S. (2000) Disturbance, patchiness, and diversity in streams. Journal of the North American Benthological Society, 19, 573-592.

Lamontagne, S., McEwan, K., Webster, I., Ford, P., Leaney, F. and Walker, G. (2004) Coorong, Lower Lakes and Murray Mouth: Knowledge gaps and knowledge needs for delivering better ecological outcomes. (CSIRO, Water for a Healthy Country National Research Flagship, Canberra, Canberra, pp. 27.

Langley, R. A., Lester, R. E., Fairweather, P. G. and Webster, I. T. (2009) Predicting the future ecological condition of the Coorong. Supplementary output from scenario modelling. Flinders University, Adelaide.

Lester, R. E. and Fairweather, P. G. (2008a) Proceedings of the CLLAMMecology Research Cluster CLLAMM Futures Workshop #2. Flinders University, Adelaide.

Lester, R. E. and Fairweather, P. G. (2008b) Review of modelling alternatives for CLLAMM Futures, CSIRO: Water for a Healthy Country National Research Flagship, Adelaide.

Lester, R. E. and Fairweather, P. G. (2009) Proceedings of the CLLAMMecology Research Cluster CLLAMM Futures Workshop #3. Flinders University, Adelaide.

Lester, R. E. and Fairweather, P. G. (in press) Modelling future conditions in the degraded semi-arid estuary of Australia's largest river using ecosystem states. Estuarine, Coastal and Shelf Science.

Lester, R. E., Fairweather, P. G. and Langley, R. A. (in prep) Ecosystem response modelling of the Coorong. In Ecosystem response modelling in the Murray Darling Basin(Ed, I. Overton and N. Saintlan) CSIRO Publishing.

Lester, R. E., Langley, R. A. and Fairweather, P. G. (2008) What are the sensitivities of key biota in the Lower Lakes to changes in salinity that may be triggered by the introduction of seawater into the Lower Lakes? A report prepared for the Murray Darling Basin Commission. Flinders University, Adelaide.

Lester, R. E., Webster, I. T., Fairweather, P. G. and Langley, R. A. (2009) Predicting the future ecological condition of the Coorong. The effect of management actions & climate change scenarios. Flinders University, Adelaide.

Murray-Darling Basin Commission (2006) The Lower Lakes, Coorong and Murray Mouth Icon Site environmental management plan 2006-2007. MDBC, Canberra, pp. 169.

Naeem, S. (2009) Ecology: Gini in the bottle. Nature, 458, 579-580.

Noell, C. J., Ye, Q., Short, D. A., Bucater, L. B. and Wellman, N. R. (2009) Fish assemblages of the Murray Mouth and Coorong region, South Australia, during an extended drought period. SARDI Aquatic Sciences, Adelaide, pp. 74.

Paton, D. C. and Rogers, D. J. (2008) 2007 winter monitoring of the southern Coorong, South Australian Department of Water Land and Biodiversity Conservation, Adelaide.

Paton, D. C., Rogers, D. J., Hill, B. M., Bailey, C. P. and Ziembicki, M. (in press) Temporal changes to spatially-stratified waterbird communities of the Coorong, South Australia: implications for the management of heterogeneous wetlands. Animal Conservation.

Petraitis, P. S. and Dudgeon, S., R. (2004) Detection of alternative stable states in marine communities. Journal of Experimental Marine Biology and Ecology, 300, 343-371.

Petraitis, P. S. and Dudgeon, S., R. (2005) Divergent succession and implications for alternative states on rocky intertidal shores. Journal of Experimental Marine Biology and Ecology, 326, 14-26.

Phillips, B. and Muller, K. (2006) Ecological character description: Coorong, Lakes Alexandrina and Albert, Wetland of International Importance. (Ed, Department for Environment and Heritage) Government of South Australia, Adelaide, pp. 336.

Plant, R. E. and Vayssieres, M. P. (2000) Combining expert system and GIS technology to implement a state-transition model of oak woodlands. Computers and Electronics in Agriculture, 27, 71-93.

Powell, J. L. (2008) Dead pool: Lake Powell, global warming and the future of water in the West, University of California Press, Berkeley.

Python Software Foundation, P. S. (2008) Python 2.3.1.

Reckhow, K. H., Arhonditsis, G. B., Kenney, M. A., Hauser, L., Tribo, J., Wu, C., Elcock, K. J., Steinberg, L. J., Stow, C. A. and McBride, S. J. (2005) A predictive approach to nutrient criteria. Environmental Science and Technology, 39, 2913-2919.

Rogers, D. J. and Paton, D. C. (2009) Changes in the distribution & abundance of *Ruppia tuberosa* in the Coorong. Adelaide University, Adelaide, pp. 22.

Rolston, A. and Dittmann, S. (2009) The distribution and abundance of macrobenthic invertebrates in the Murray Mouth and Coorong. Flinders University, Adelaide.

Steinberg, D. and Golovnya, M. (2007) CART 6.0 User's guide, Salford Systems, San Diego.

Suding, K. N. and Hobbs, R. J. (2009) Threshold models in restoration and conservation: a developing framework. Trends in Ecology and Evolution, 24, 271-279.

Sutherland, G. D. (2006) Predicting the ecological consequences of environmental change: a review of the methods. Journal of Applied Ecology 43, 599-616.

Therneau, T. M. and Atkinson, B. (2009) Package 'rpart'.

van Nes, E. H. and Scheffer, M. (2005) Implications of spatial heterogeneity for catastrophic regime shifts in ecosystems. Ecology, 86, 1797-1807.

Webster, I. T. (2005) An overview of the hydrodynamics of the Coorong and Murray Mouth. CSIRO Water for a Healthy Country National Research Flagship, Canberra.

Webster, I. T. (2006) Hydrodynamic modelling of the Coorong, Water for a Healthy Country National Research Flagship, CSIRO, Canberra.

Westoby, M., Walker, B. H. and Noy-Meir, I. (1989) Opportunistic management for rangelands not at equilibrium. Journal of Range Management, 42, 266-274.

Wilkinson, S. R., Naeth, M. A. and Schmiegelow, F. K. A. (2005) Tropical forest restoration within Galapagos National Park: Application of a state-transition model. Ecology and Society, 10, 28.

Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W. and Boon, N. (2009) Initial community evenness favours functionality under selective stress. Nature, 458, 623-626.

Wolanski, E. (1987) An evaporation-driven salinity maximum zone in Australian tropical estuaries. Estuarine, Coastal and Shelf Science, 22, 415-424.

Young, W. J., Scott, A. C., Cuddy, S. M. and Rennie, B. A. (2003) Murray flow assessment tool: A technical description. Client report, 2003. CSIRO Land and Water, Canberra.

Zar, J. H. (1999) Biostatistical analysis, PrenticelHall, New Jersey.

Zweig, C. L. and Kitchnes, W. M. (in press) Multi-state succession in wetlands: a novel use of state and transition models. Ecology.

Appendix A – Road map to the outputs from the CLLAMM Futures theme

The following diagrams illustrate the various outputs from the CLLAMM Futures theme. They show the sequence in which they will be or have been prepared, their current stage of completion and the links between the various outputs (Figure A.1). Figure A.2 also shows the intended audience and purpose of each output, to highlight the range of target audiences covered.

Figure A.1 Road-map of CLLAMM Futures outputs

* The CLLAMM Futures final reporting is made up of three separate reports, including this document, a scenario analysis report and a supplementary report.

Key
- Purpose
- Audience
- Target publication
- Date published (where applicable)

Workshop Report 1
- Synthesis of discussion from Workshop 1
- CLLAMM researchers & stakeholders
- CSIRO website
- May 2007

Workshop Report 2
- Synthesis of discussion from Workshop 2
- CLLAMM researchers & stakeholders
- CSIRO website
- April 2008

Review of modelling alternatives
- Summaries & assesses potential modelling options
- CLLAMM researchers, scientific community
- CSIRO website, hardcopy
- May 2008

Introducing the Coorong and its ecosystem states
- Introduce the state of the Coorong & the approach of ecosystem states
- International scientific community
- Special issue of *Estuarine, Coastal and Shelf Science* (from ECSA 44)

ERM in the Coorong
- Describe results of ecosystem state modelling
- Ecologists & managers in MDB
- Ecological response modelling in the Murray-Darling Basin book

Workshop Report 3
- Synthesis of discussion from Workshop 3
- Invited participants, international & interstate ecologists
- CSIRO website

CLLAMM Futures Final report
- Detailed methodology & development, the models themselves, all scenario analyses & management implications
- CLLAMM researchers & managers
- CSIRO website, hardcopy

Predictive model development and verification
- Justification of model, scenario analyses, comparison of few scenarios
- Modelling scientific community
- Modelling journal

Meta-analysis with other estuaries
- Place Coorong options in southern Australian perspective, provide benchmarks for recovery
- Ecological scientific community (mainly Australian)
- Australian ecological journal

Results of scenario analyses
- Ecological assessment of selected scenarios
- Ecological scientific community
- Applied ecological journal

Other manuscripts to be determined
- Specific subsets of data (with other CLLAMM researchers)
- International scientific community
- Applied journals

Figure A.2: Purpose and intended audience of CLLAMM Futures outputs

Appendix B – Summary of modelling and analysis techniques investigated

The following table summarises the techniques explored in the modelling review that was undertaken (Lester and Fairweather, 2008b), listing their advantages, disadvantages and their applicability for CLLAMM Futures, as determined at the time.

	Advantages	Disadvantages	Applicability to Futures
Hysteresis modelling	Leads to greater understanding of state change and system predictability	Describes an outcome, rather than a particular technique	Appropriate data to model hysteresis is unlikely to be available
Alternative stable states	Leads to greater predictability of future states, as systems with alternative stable states show memory of previous states	Difficult to adequately demonstrate true alternative stable states without experimentation	The alternative state concept is useful, although it is unlikely that stability will be demonstrated
Bifurcation plots from chaos theory	Able to describe the differences in parameter values across a threshold value Understanding one bifurcation can assist in predicting others so can make systems more predictable	Requires data on the behaviour of systems both in the decline and recovery phase Largely limited to use in deterministic models	Unlikely to be useful given Future's focus on statistical, rather than deterministic modelling
State and transition modelling	Able to combine data from different sources to describe states and/or transitions Does not assume equilibrium in a system Can be linked to GIS to become spatially explicit	Have limited ability to account for gradual change Are not able to incorporate evolutionary change	Could be used as the structure for the various hypothesised ecosystem states, combining output from other techniques to define the various states and transitions

Table B.1. Summary of the relative advantages, disadvantages and applicability of the techniques investigated

	Advantages	Disadvantages	Applicability to Futures
Classification and regression trees	Are able to use a combination of categorical and continuous variables Are well suited to ecological data due to a lack of assumptions about distributions Have easy-to-understand graphical output	A lack of consensus in algorithms mean that different programs can result in different output trees Tend to produce overly-complex trees unless pruned appropriately	Could be used to parameterise the hypothesised alternative ecosystem states Able to identify threshold points between states
Multiple regression trees	Extends CART analyses to situations that have multiple response variables Able to assess high-order, complex interactions between predictor variables Robust to patchy ecological data	As per CART, is prone to over-fitting Can be sensitive to outliers and be biased towards variables with more potential split points.	Is likely to be of use in predicting species composition under tested scenarios May be able to validate hypothesised ecosystem states
Boosted regression trees	Greater predictive accuracy than CART because a combination of CART models are fitted to progressively re-weighted data Similar advantages to other tree-based analyses, like the ability to incorporate both categorical and continuous variables	The output is complex, with up to hundreds of simple trees combined in one analysis with individual trees unable to be examined individually Prone to over-fitting unless the output is carefully pruned	Unlikely to be useful given the "black box" nature of the output that will preclude the results from being incorporated into other analyse

Table B.1 cont. Summary of the relative advantages, disadvantages and applicability of the techniques investigated

	Advantages	Disadvantages	Applicability to Futures
Multivariate adaptive regression splines	Able to detect global and linear structures within data sets Combination of regression equations leads to greater predictive power than traditional regression equations. Found to be effective in combination with other regression analysis techniques like GAM and GLM.	Some contention regarding utility in ecological situations. The ability to incorporate output into other analyses is unclear and may vary between analytical packages. Has the potential to be over-influenced by local processes, making findings unstable	The technique may be useful, but verification with CLLAMM Futures data is needed to confirm, given varying findings in the scientific literature.
Structural equation modelling	Tests hypothesised relationships between variables Able to propagate error and uncertainty through the model Able to incorporate theoretical constructs, as well as measured variables	Can be difficult to include spatial and temporal scales into the model Can be difficult to assess causality from co-variations and correlations	Could be used to quantify the conceptual models used by managers of the CLLAMM region
Bayesian belief networks	Represents a system as a series of interactions Able to combine a variety of data and analysis types New information can be incorporated as it becomes available	Require a full probability structure to be specified Poor at incorporating spatial or temporal variability Continuous variables are not treated neatly Historically, tend to rely overly on expert opinion	Could be used as the platform to tie various other analyses together

Table B.1 cont. Summary of the relative advantages, disadvantages and applicability of the techniques investigated

	Advantages	Disadvantages	Applicability to Futures
Gaussian error propagation	Enables detailed study of error sources within the model Computationally less expensive than techniques like Monte Carlo simulations	Output variables must be expressed as functions of input variables	CLLAMM models will not be able to be expressed as a single mathematical function
Individual-based modelling	Scale consistent with evolutionary adaptation Findings are likely to be predictive because they are based on theory and are not empirical	Highly data intensive Involves high-level programming skills Often involves studies over decades Can be difficult to communicate findings	Appropriate data will not be available for CLLAMM
Levels of evidence	Combines multiple studies to strengthen the evidence for or against a hypothesis Rigorous method of hypothesis falsification over several steps Identifies areas for further research	Time-consuming technique	Unlikely, but may be useful in identifying knowledge gaps for possible future work in the region

Table B.1 cont. Summary of the relative advantages, disadvantages and applicability of the techniques investigated

Appendix C – Species list from 1984/85

The following table gives the species that were recorded in 1984 and/or 1985 in data sets that were used in the evaluation of the model described here. The sources of the data are given above in the text.

Species	Year 1984	1985	1999-2007
Macrophytes			
Ruppia tuberosa	X	X	X
Ruppia megacarpa	X	X	
Lepilaena sp.	X	X	
Zostera muelleri	X	X	
Macroinvertebrates			
Crustaceans			
Amarinus lacustris*			
Macrobracium intermedium	X	X	
Melita zeylanica	X	X	
Paracorophium sp.*	X	X	
Megamphopus sp.	X	X	
Polychaetes			
Ceratonereis pseudoerythraensis	X	X	X
Nephtys australiensis	X	X	X
Prionospio sp.	X	X	
Ficopomatus enigmaticus*	X	X	
Boccardia sp.	X	X	
Capitella capitata	X	X	X
Capitellides	X	X	X
Fabiciinae sp.	X	X	
Phyllodoce novohollandiae			
Molluscs			
Hydrobia buccinoides	X	X	X
Notospisula trigonella	X	X	
Arthritica semen	X	X	X
Tellina sp.			
Tatea sp.			
Salinator fragilis	X	X	X
Insects			
Chironomid sp.	X	X	X
Ephydrella sp.	X	X	
Fish			
Australian Salmon (*Arripis truttaceus)*	X	X	X
Black Bream *(Acanthopgarus butcheri)*	X	X	X
Bony Bream *(Nematalosa erebi)*	X	X	X
Mulloway *(Argyrosomus hololepidotus)*	X	X	X
Yellow-eye Mullet *(Aldrichetta forsteri)*	X	X	X
Jumping Mullet *(Liza argentea)*	X	X	X

Table C.1. Species list for 1984 and 1985

Note: Species annotated with an asterisk were not in the training data set but were observed in the system during CLLAMMecology by researchers involved in the Research Cluster.

Species	Year		
	1984	**1985**	**1999-2007**
Callop *(Macquaria embigua)*	X	X	X
Redfin Perch *(Perca fluviatilis)*	X	X	X
Greenback Flounder *(Rhombosolea tapirina)*	X	X	X
European Carp *(Cyprinus carpio)*	X	X	X
Bronze Whaler Shark *(Carcharhinus brachyurus)*	X	X	X
Gummy Shark *(Mustelus antarcticus)*	X	X	X

Birds

Species	1984	1985	1999-2007
Australian Pelican *(Pelecanus conspicillatus)*	X	X	X
Little-pied Cormorant *(Phalacrocorax melanoleucos)*	X	X	X
Pied Cormorant *(Phalacrocorax varius)*	X	X	X
Little Black Cormorant *(Phalacrocorax sulcirostris)*	X	X	X
Great Cormorant *(Phalacrocorax carbo)*	X	X	X
Black-faced Cormorant *(Phalacrocorax fuscescens)*		X	X
Black Swan *(Cygnus atratus)*	X	X	X
Australian Shelduck *(Tadorna tadornoides)*	X	X	X
Musk Duck *(Biziura lobata)*	X	X	X
Pacific Black Duck *(Anas superciliosa)*	X	X	X
Chestnut Teal *(Anas castanea)*	X	X	X
Grey Teal *(Anas gracilis)*	X	X	X
White-faced Heron *(Ardea novaehollandiae)*	X	X	X
Pied Oystercatcher *(Haematopus longirostris)*		X	X
Red-capped Plover *(Charadrius ruficapillus)*	X	X	X
Masked Lapwing *(Vanellus miles)*	X	X	X
Black-winged Stilt *(Himantopus himantopus)*	X	X	X
Banded Stilt *(Cladorhynchus leucocephalus)*	X	X	X
Red-necked Avocet *(Recurvirostra novaehollandiae)*	X	X	X
Fairy Tern *(Sterna nereis)*	X	X	X
Crested Tern *(Sterna bergii)*	X	X	X
Caspian Tern *(Hydropogne caspia)*	X	X	X
Whiskered Tern *(Childonias hybridus)*	X	X	X
Silver Gull *(Larus novaehollandiae)*	X	X	X
Sharp-tailed Sandpiper *(Calidris acuminata)*	X	X	X
Red-necked Stint *(Calidris ruficollis)*	X	X	X
Common Sandpiper *(Actitis hypoleucos)*	X	X	X
Common Greenshank *(Tringa nebularia)*	X	X	X
Great Crested Grebe *(Podiceps cristatus)*	X	X	X
Hoary-headed Grebe *(Poliocephalus poliocephalus)*	X	X	X

Table C.1 cont. Species list for 1984 and 1985