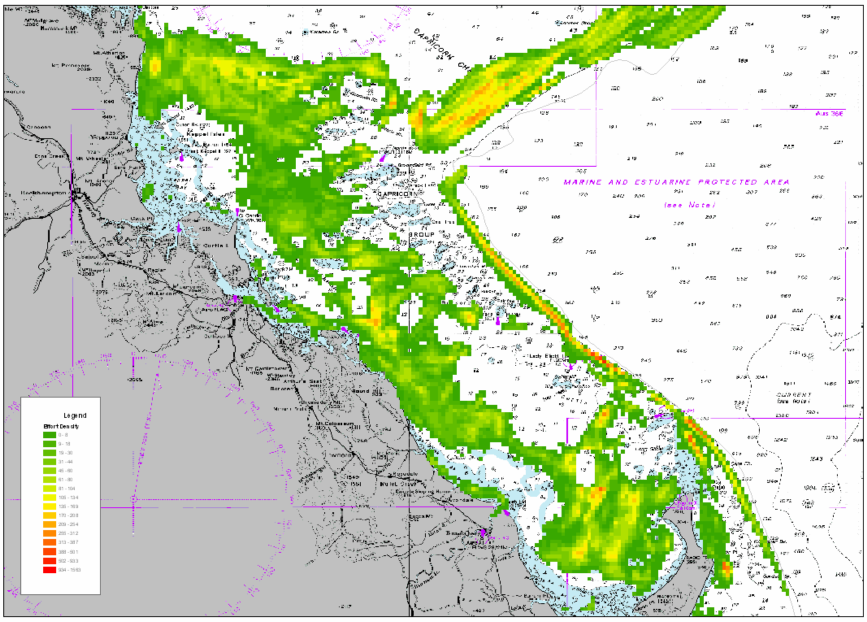
Queensland **the Smart State**

Innovative stock assessment and effort mapping using VMS and electronic logbooks

FRDC Project 2002/056 Final Report
November 2007



Principal Investigator:
Neil Gribble

Authors:
Norm Good
David Peel
Mai Tanimoto
Rick Officer
Neil Gribble

Australian Government
Fisheries Research and Development Corporation

Queensland Government
Department of Primary Industries and Fisheries

Innovative stock assessment and effort mapping using VMS and electronic logbooks

Principal Investigator
Neil Gribble

Authors
Norm Good[1], David Peel[2],
Mai Tanimoto[3], Rick Officer[3], Neil Gribble[4]

Queensland Government
Department of Primary Industries and Fisheries

Australian Government
Fisheries Research and Development Corporation

Project No. 2002/056

[1] CSIRO Mathematical and Information Sciences, Brisbane, QLD 4000
[2] CSIRO Mathematical and Information Sciences, Hobart, Tasmania, 7001
[3] Department of Primary Industries and Fisheries, Southern Fisheries Centre, Deception Bay, QLD 4508
[4] Department of Primary Industries and Fisheries, Northern Fisheries Centre, Cairns, QLD 4870

PR07-3229
ISBN: 978 0 7345 0379 4

© The State of Queensland, Department of Primary Industries and Fisheries, Fisheries Research and Development Corporation, 2007

This work is copyright. Except as permitted under the *Copyright Act 1968 (Cth)*, no part of this publication may be reproduced by any process, electronic or otherwise, without the specific written permission of the copyright owners. Neither may information be stored electronically in any form whatsoever without such permission.

This work should be cited as:
Good, N., Peel, D., Tanimoto, M., Officer, R. and Gribble, N. (2007). Innovative stock assessment and effort mapping using VMS and electronic logbooks. Final report on FRDC project 2002/056. Department of Primary Industries and Fisheries, Brisbane.

The Fisheries Research and Development Corporation plans, invests in and manages fisheries research and development throughout Australia. It is a statutory authority within the portfolio of the federal Minister for Agriculture, Fisheries and Forestry, jointly funded by the Australian Government and the fishing industry.

Inquiries should be addressed to:
Intellectual Property and Commercialisation Unit
Department of Primary Industries and Fisheries
GPO Box 46
Brisbane QLD 4001

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AFMA	Australian Fisheries Management Authority
AIC	Akaike's Information Criterion
AMSA	Australian Marine Science Association
BIC	Bayesian Information Criterion
CFISH	Commercial Fisheries Information System
CPUE	catch per unit effort
DEH	Australian Government Department of the Environment and Heritage
ECERS	Electronic Catch and Effort Reporting System
EKP	Eastern King Prawn
EM algorithm	Expectation-Maximisation algorithm
ESD	Ecologically Sustainable Development
FARM	Fisheries Activity Research Modelling
GBRMPA	Great Barrier Reef Marine Park Authority
GVP	Gross Value of Production
HMM	Hidden Markov Model
ICF	Intrinsic correlation function
IOI	International Oceans Institute
Logbook unit	Assessment & Monitoring unit
LPUE	landings per unit effort
MPA	Marine Protected Area
MaxEnt	Maximum Entropy
MSE	mean squared error
nm	nautical mile
QDPI&F	Queensland Department of Primary Industries and Fisheries
QSIA	Queensland Seafood Industry Association
RAP	Representative Areas Program
REML	Residual Maximum Likelihood
SRAs	Scallop Replenishment Areas
Trawl MAC	Trawl Fishery Management Advisory Committee
VMS	Vessel Monitoring Systems
VMS unit	Fisheries Vessel & Quota Monitoring Services Unit

2002/056 Innovative stock assessment and effort mapping using VMS and electronic logbooks

PRINCIPAL INVESTIGATOR: Dr N. Gribble
ADDRESS: Queensland Department of Primary Industries and Fisheries (QDPI&F)
Northern Fisheries Centre
PO Box 5396
Portsmith, Queensland 4870
Telephone: 07 4035 0128 Fax: 07 4035 1401

OBJECTIVES:

1. Review applications and potential of VMS mapping and OceanFARM software and related approaches.
2. Develop trawl track and trawl signature definitions for each fishery sector to use with TerraVision software.
3. Map the spatial and temporal intensity of fishing effort in each trawl sector, and estimate the distribution and extent of trawled and untrawled areas.
4. Map resource density indices for each fishing sector.
5. Use these methods to recommend (and achieve implementation of) improved Trawl Fishery Review Events, and develop improved stock assessment approaches for scallops, eastern king prawns (EKP) and tiger/endeavour prawns.

OUTCOMES ACHIEVED TO DATE

- The most important outcome from the project has been the adoption of the project's VMS mapping and trawl signature recognition algorithms/software 'TrackMapper' by management (i.e. QDPI&F VMS unit).
- An equally important outcome has been the acceptance by the fishing industry of the importance of VMS data in fishing effort mapping and resource assessment (demonstrated by the 2005 Queensland Seafood R&D Award from the fishing industry to the VMS project).
- Another major outcome was the application of the methods developed by the project to produce maps detailing the amount of Gross Value of Production (GVP) lost as a result of the introduction of the Representative Areas Program. This work was done for the Australian Government Department of the Environment and Heritage (DEH) to aid in developing a structural readjustment package to compensate fishers. The total value of the package may go as high as A$100 million.
- A further outcome has been the enhancement of the basic concepts and computer algorithms that will drive the future development of fisheries

resource assessment using high-resolution VMS data and electronic logbooks.

Note: Confidentiality of fishers VMS data was negotiated at the beginning of the project and was subject to a mutually agreed ‘Confidentiality Deed’ (see Appendix C). All presentations of data in this report and in all subsequent publications are in compliance with this deed.

1 NON-TECHNICAL SUMMARY

A satellite-based Vessel Monitoring System (VMS) provides real-time position information of fishing boats at intervals depending on the interrogation (polling) frequency. As such, VMS data represent an improvement in the quality of fisheries spatial data that will profoundly affect the way fisheries statistics are used for fish population modelling, and ultimately, fishery management. This applies not only to Queensland fisheries but also to other state and international fisheries as VMS is progressively adopted worldwide.

This project used VMS position information and logbook catch records to map and statistically model the spatial distribution of Queensland otter trawl fisheries, and estimate density indices for selected fisheries. These high-resolution maps and indices have been used in several management reviews and subsequently adopted within improved management procedures.

The overarching aim of the project was to enable better trawl fishery management by providing both better quality and global information, and leading to changes in management arrangements. Specifically, the project set out to:

1) Review applications and potential of VMS mapping and resource modelling software and related approaches.
2) Develop trawl track and trawl signature definitions for VMS position data for each fishery sector.
3) Map the spatial and temporal intensity of fishing effort in each trawl sector, and estimate the distribution and extent of trawled and untrawled areas.
4) Map resource density indices (the underlying prawn abundance) for each fishing sector, using the eastern king prawn (EKP) (*Penaeus plebejus)* fishery as a case study.
5) Identify potential improvements to stock assessment inputs that could be provided through the use of VMS data.

Since most VMS implementations record vessel location at set time intervals with no regard to vessel activity, a methodology is required to determine which position data correspond to fishing activity, i.e. to identify strings of position data that are characteristic of trawling or ‘trawl signatures’. Separate trawl signatures were developed for the EKP, tiger/endeavour prawns (*Penaeus esculentus*, *P. semisulcatus*, and *P. monodon*/*Metapenaeus endeavouri* and *M. ensis*), scallop (*Amusium japonicum* and *A. pleuronectes*) and banana prawn (*Fenneropenaeus merguiensis*) fisheries using a sequence of decision rules to filter VMS hourly polling data.

Speed Rule(s)
The simple observation that vessels travel much slower while trawling than when cruising allows for discrimination based simply on the speed of the vessel. Speed data are either included in the information transmitted by the vessel when interrogated by VMS, or derived from the time and distance between sequential VMS positions.

A ‘mixture model’ was used as a statistical approach to classify the speed data into two groups: trawling and steaming. A two-component normal mixture model was applied to the historical VMS speed data from each fishery. The critical point or boundary between the trawling and steaming groups was taken as the upper trawl signature speed.

Alternatively, a probabilistic approach based on Hidden Markov Models (HMM) can be used to determine vessel activity from VMS speed data. The use of a probabilistic framework rather than an outright classification allows a better indication of the uncertainty in the allocation. The HMM provides a natural framework for the problem and by definition models the intrinsic temporal correlation of the data. This report describes the general approach that was developed and presents an example of this approach applied to Queensland’s East Coast trawl fisheries. Finally, the report presents the results of a validation and error quantification experiment for the HMM approach.

Time of day Rule
The scallop fishery and most prawn fisheries are predominantly conducted during the night. An astronomical algorithm which tracks the exact dusk and dawn times was incorporated into database queries to exclude daytime polls. [Note: All polls over a 24-hour period were selected for the banana prawn and stout whiting (*Silago robusta*) fisheries, as both fisheries can be day and night operations].

Ports and anchorages Rule
After the application of speed and time of day rules there was an obvious problem of effort in ports and anchorages. In these cases a vessel may have been steaming for part of an hour, while either entering or leaving port. This can result in a calculated speed within the trawl speed range being produced. Therefore, a decision rule to identify a port or anchorage and remove the resulting polls was added.

Subsequent maps showing the intensity and distribution of fishing effort were produced for the EKP, tiger/endeavour prawns, scallop and all otter trawl fisheries along the entire eastern coastline.

To produce comparable maps of spatial patterns in abundance of the fished species, it was necessary to assign the reported catch to the high-resolution maps of fishing effort that the project produced from the VMS data. Maximum entropy analysis of VMS-derived effort data and commercial catch logbook data was used to infer the stock distribution of the EKP fishery. This fishery has deep (more than 90 metres) and shallow (less than 90 metres) sectors that were analysed separately. The incorporation of an intrinsic correlation function into the maximum entropy model resulted in the ability to investigate prawn stock distribution at spatial scales reduced to two nautical mile (nm) square grids in both sectors. Three- and seven-day subsets of data were used to model stock distribution and density for the deep and shallow sectors, respectively. Using these subsets ensured that a maximum entropy model fit could be produced whilst the likelihood of ‘blurring’ the image as a result of stock movement was reduced. Maximum entropy predicted stock distribution reasonably well, although it consistently overestimated actual density at high values.

While showing great potential, the method highlighted that effort and catch data must be at similar fine spatial resolution to improve the accuracy of density indices. There was a problem of temporal blurring if the effort data were collected over a short time period (duration of trawl) but the catch records were integrated over a longer period (a 'night's' fishing, effectively 24 hours). Even with shot-by-shot catch records, this problem may remain.

The VMS mapping algorithms and software developed by the project are now standard operating procedures for the Fisheries Vessel & Quota Monitoring Services Unit (VMS unit) of Queensland Department of Primary Industries and Fisheries (QDPI&F). Maps generated by this project and by the QDPI&F VMS unit have been used in a number of fishery reviews and stock assessments, for example:

- To describe areas fished and areas now protected relative to areas previously fished, in negotiations for compensation of trawl fishers in industry restructuring
- Fine detail mapping of VMS information was incorporated into the current standardisation model for the EKP fishery and used to evaluate CPUE-related reference points
- Fine detail mapping of VMS information was incorporated into stock assessment of the Torres Strait Tiger Prawn Fishery (*Penaeus esculentus*). This was a major review event prompted by legislative change to the management plan.

To improve parameter estimates for stock assessment, the project explored estimating the 'catchability' q using VMS-derived data. This parameter relates CPUE indices of abundance to the underlying absolute abundance of a stock. Depletion analyses were used as they produce direct estimates of q.

Whilst showing promise, the use of VMS spatial information in depletion analyses was not straightforward. As with the Maximum Entropy analysis, the depletion study highlighted that the scales and resolution of effort and catch data must be similar to get an accurate depletion estimate.

For all three fisheries explored, fishing effort was found to be spatially aggregated:

- Analysis of the tiger/endeavour prawns fishery VMS data suggested that targeting occurred in areas of high CPUE; interpreted as fishers targeting aggregations of prawns.
- Analysis of the scallop and EKP fisheries, in contrast, showed that although effort was aggregated it was not related to areas of high CPUE; interpreted as fishers following spatial patterns determined by external processes which might include management closures, fuel prices, and cost-benefit business decisions.

Future calculation of CPUE indices should involve the application of appropriate spatial weightings to correct for the concentration of fishing effort within the fishery area, and for the targeting of areas with higher catch rates.

This project experienced recurring difficulties due to data deficiencies related to the assignment of activities and catch to vessel tracks. Improved definition of vessel activity was considered a tractable problem. Novel technological solutions may diminish the need for estimation procedures that define vessel activity through the direct recording of data that accurately characterises vessel activity. Technological solutions could also be used to collect high-resolution catch and effort data that may empirically validate the precision provided by the low-resolution commercial data that are currently available.

KEYWORDS: Vessel Monitoring System, Maximum Entropy, trawl definition, trawl signatures, VMS mapping, mixture model, Hidden Markov Model

2 BACKGROUND

2.1 VMS and electronic logbooks

VMS provides real-time locations of fishing boats, at intervals depending on the polling frequency. VMS data represent an improvement in the quality of fisheries spatial data that will profoundly affect the way fisheries statistics are used for fish population modelling, and ultimately fishery management. This applies not only to Queensland fisheries but also to other state and international fisheries, as VMS is progressively adopted worldwide.

Electronic logbook systems, such as Queensland's Electronic Catch and Effort Reporting System (ECERS), or AFMA's electronic logbook system, allow boats to log their compulsory daily catch records electronically. This has a number of advantages over alternative manual systems, including reduced error and timely receipt of data by fisheries management.

Up-to-date datasets, together with precise spatial data from VMS, offer great potential to improve stock assessment and management systems. Queensland has the first Australian fisheries, and among the first in the world, to install both VMS and ECERS. This is an opportunity to be at the international forefront of developing fishery information and management systems.

2.2 TerraSystems software: OceanFARM and TerraVision

OceanFARM (Fisheries Activity Research Modelling) is a recently developed system for fisheries modelling with spatial data from VMS and catch data from electronic or manual logbooks. It provides tools to model temporal and spatial fishery dynamics, and estimate density indices for each resource, based on the distribution of catch rates in space and time. (See 6.6.1.)

TerraVision is contributing $1.44m towards the OceanFARM project, and a Commonwealth Government R&D Start Grant is contributing $0.96m (TerraVision's contribution is being audited on behalf of AusIndustry by Hayes Knight GTO). The OceanFARM project is also contributing $135,000 to the QDPI&F in return for access to ECERS data. The VMS and ECERS in Queensland have been developed as a joint project between the QDPI&F and TerraVision.

2.3 Ecological sustainability and management objectives

Ecologically Sustainable Development (ESD) is required in Queensland under the *Fisheries Act 1994* and under the Commonwealth *Environment Protection and Biodiversity Conservation Act 1999*. Assessment and management of risks to sustainability are urgent priorities for all Queensland-based fisheries servicing domestic or export markets. The Queensland East Coast Trawl Fishery is of particular interest as it operates largely within the Great Barrier Reef World Heritage Area. It therefore has a critical need to quantify and manage risk to the ecological sustainability of its fishing grounds. The fishery yields product worth about A$140 p.a. and also has a profound economic downstream effect on local fishing communities and support infrastructure along the 2000 km coastline.

Management of the eastern king prawn (EKP), scallop, and tiger prawn fisheries uses reference points and Fishery Review Events based on catch per unit effort (CPUE). However, it is well established that CPUE on a broad spatial scale can be a poor index of abundance, since targeting of high-density patches and communication between skippers promotes hyperstability in catch rates. Spatial density indices of the type given by the OceanFARM system may provide much better indices of abundance, and hence much more reliable reference points.

Stock assessment models were developed for the EKP, scallop, and tiger prawn fisheries as part of the trawl effort standardisation and target reference points project (FRDC Project # 1999/120: O'Neill, Courtney *et al.* 2005). Reference points and review events based on spatial density indices will require further development of these models. In addition, new types of information will permit further enhancement to increase the power of these models.

2.4 Modelling issues: trawl signature, polling frequency, and uncertainty

To model effort and catch distribution we must estimate when a boat is trawling, based on VMS data. This is known as the 'trawl signature', and must be developed separately for each fishery sector, or even for areas and times within a sector. Indices of density and distribution are imprecise, and this uncertainty must also be estimated. Uncertainty will vary between sectors and with the polling frequency. The cost associated with any increase in polling frequency must be justified by increased return to the fishery. The current polling frequency of one per hour has provided far better information about resource and effort distribution than was previously available.

2.5 Data standards: VMS

This project complies with QDPI&F's open policy regarding VMS data structures, in that the QDPI&F VMS data are being sent on to Great Barrier Reef Marine Park Authority (GBRMPA), Australian Marine Science Association (AMSA), and WEBVision in real time. Coastwatch have also declared their interest. All these organisations use TerraVision, as do Western Australian and Tasmanian Fisheries. NSW Fisheries have expressed interest in going online as well. AFMA's VMS data specification is defined in a commercial off-the-shelf Smartrac Oracle database. Both Smartrac and TerraVision systems use vessel number, vessel name, latitude (decimal), longitude (decimal), report date, report time, measured speed and measured direction as fundamental components of their position database structure.

2.6 Data standards: catch and effort

Each fishery data entry form has its own table of information, different to other tables of information for other reporting regimes. Thus, unless there are two identical fisheries operating in different jurisdictions with identical reporting requirements, standards can only be broadbrush.

OceanFARM is not tied to any particular logbook system. For example, OceanFARM can use data from the ECERS and the manually sourced logbook system (CFISH). OceanFARM can also be readily adapted to AFMA's electronic logbook system. The map that accompanies this application was produced using CFISH data.

2.7 Data standards: modelling

Australia has no standard for model data. However, TerraVision through its OceanFARM development has developed a specification for its FARM data set which incorporates raw and processed modelling data. The FARM data set draws on the developer's 20 years of experience (MineMap was founded in 1981) in spatial modelling and is being offered to research as an open platform and as a tool on which research can build its own modelling systems.

2.8 2006 Update on background

(a) The need to investigate increased polling rates in order to improve spatial resolution of effort maps was partially satisfied by the QDPI&F's agreeing to increase the VMS polling rate from four per day up to one per hour. This rate increase was initially trialled for 12 months then maintained for at least the life of the project. All VMS data were also archived for use with this project, as part of the QDPI&F contribution.

(b) Although the resolution of the VMS polling was increased, ECERS suffered from technical and political problems. At the beginning of project there was an increasing participation rate by trawler operators, however the political process for the review of the Trawl Management Plan, Representative Areas Program (RAP), and the subsequent industry restructuring/compensation, reversed this trend. Instead of the anticipated take-up of ECERS by the majority of the fleet only very few operators have persevered (2–5 boats).

(c) Because the 'real-time' ECERS data did not eventuate, TerraVision's development of OceanFARM stalled. Our collaboration with TerraVision on those elements of OceanFARM that required the higher resolution catch information also stalled. To meet project commitments and milestones more emphasis was placed on making maximum use of 'paper logbook' records and validation using GPS records from individual trawlers. Again the latter collaboration suffered because of the deteriorating political environment.

These adjustments are consistent with the Methods as stated in the project proposal, i.e. 'The project is designed as a desktop exercise using existing data, outcomes from current research projects, and additional VMS, ECERS and logbook data as they become available'.

3 NEED

3.1 Need for trawl mapping

Information on where trawling does and does not occur is needed by fishery managers, industry, GBRMPA and others to inform debate and decision making for the trawl fishery. By June 2002, VMS have recorded all Queensland trawl effort (except the Moreton Bay fishery) every hour for 18 months. These data can be used to map the distribution and intensity of trawling better than ever before. These maps were required by July 2003 for implementation of the Queensland Trawl Plan. Such maps were also needed to model the ecological effects of trawling, since untrawled areas may provide refuge for some vulnerable bycatch species. Such maps will also help to assess the required 40 per cent reduction in bycatch.

3.2 Need to develop stock assessment and management for ESD

The trawl fishery Management Advisory Committee (Trawl Mac) has named stock assessment and Review Events as their top research priorities, and VMS research as a high priority. There is a need to improve abundance indices, currently based on CPUE from trawl shots defined as square CFISH grids (30 minutes by 30 minutes). This is unrealistic and can lead to significant errors in stock assessment. There is also a need to investigate the way targeting and depletion of aggregations potentially interact with economic factors to affect CPUE.

We can meet these needs using effort and density indices at fine spatial and temporal scales, by using the functionality of newly developed commercial software to develop our modelling systems. Matrices of stock abundance in space and time can be mapped or used in stock assessment models. A major area of research need with the OceanFARM software is user definition of trawl signature and catch distribution functions, which differ between sectors of the trawl fishery.

The functionality must be integrated into the overall management and assessment strategy for each fishing sector. There is potential to substantially improve the reliability of stock assessments.

4 OBJECTIVES

The overarching objective is to enable better trawl fishery management by providing both better quality and new types of information, and by achieving changes in management arrangements.

1. Review applications and potential of VMS mapping and OceanFARM software and related approaches.
2. Develop trawl track and trawl signature definitions for each fishery sector, to use with TerraVision software.
3. Map the spatial and temporal intensity of fishing effort in each trawl sector, and estimate the distribution and extent of trawled and untrawled areas.
4. Map resource density indices for each fishing sector.
5. * Use these methods to recommend (and achieve implementation of) improved Trawl Fishery Review Events, and develop improved stock assessment approaches for scallops, eastern king prawns and tiger prawns (*modified in 2006, see Variations).

5 VARIATIONS

The sequential loss of all three original biologists on the project (Simon Hoyle to South Pacific Commission, New Caledonia; David Peel to CSIRO, Hobart; Norm Good to CSIRO, Brisbane) required that the VMS 2002/056 project team review the project schedule at the start of 2006 to determine the most effective strategy to consolidate the gains achieved so far by the project. Over the three-year period of the project Dr Neil Gribble had taken over as Principal Investigator; Norm Good had taken on both project biologist and programmer roles; Mai Tanimoto, who worked alongside Norm Good, took over Norm's role when he left, and assisted by Dr Rick Officer, worked on the project report for the last six months through to its completion.

In parallel with staff re-deployment, we also negotiated a change of wording to the June 2006 milestone and Objective 5 to:

* ‘Assess the Trawl Fishery Management Plan Review Events and other reference points given the improved spatial definition of the data. Develop

improved data inputs for stock assessments, using the EKP stock as a case study.'

The revised milestone wording and change in the underlying objective (Objective 5) is in line with the FRDC position on not funding stock assessments *per se*, but still provides a summary example of what the project aimed to achieve; i.e. improved parameter estimates based on satellite VMS data/technology. The original wording of Objective 5 and its milestones were too ambitious (i.e. to develop three new stock assessment models, complete with management strategy evaluations for each fishery) and, given events beyond control of the project, we believe were outside the scope of the main objectives.

As Principal Investigator, Dr Neil Gribble communicated with the FRDC (Crispian Ashby) by telephone (1 June 2006) and by email (21 June 2006) about the need for these changes and subsequently received agreement on the above variation.

6 REVIEW OF VMS-RELATED STUDIES

David Peel, Norm Good, and Neil Gribble

6.1 Introduction

This review formed part of the second milestone for the project. The relevant objective was to 'review the utility and applications of the OceanFARM software and related approaches'. This objective was met with a review of literature and software available at that time and relevant to this project, and has been updated for the final report.

The first part of the review is broken up into sections that correspond to the main uses of VMS data covered in the literature. This is followed by a review of methods in the literature that are relevant to this project. In particular the Bayesian Maximum Entropy algorithm shall be described and discussed. The review of the OceanFARM software as specified by the project milestone is given in the context of a software review in Section 6.6.1.

6.2 Using VMS data for more than compliance

Generally VMS has been implemented as a compliance and monitoring tool. Only recently has the mass of detailed information that VMS produces been considered for other uses, such as effort or resource-intensity mapping. A few papers have conducted either small pilot studies or desktop simulations to examine the feasibility of using the VMS system for more than compliance. However, no full studies with complete fleet coverage, as conducted in this project, could be found apart from the work of Deng *et al.* (2005). This may have been because, as stated by Nishida and Booth (2001), much of the work of this type is done in-house and not suitable for publication in peer-reviewed literature, although this is changing.

Hall *et al.* (1999) examined, through a simulation study, the viability of using VMS data collected in the Northern Australian Prawn Fishery. Their main focus was to determine if the polling frequency used by the fisheries VMS was sufficiently high. As well as using the actual VMS data, Hall *et al.* (1999) used GPS data to simulate VMS data at varying polling frequencies. They found that at the relatively low polling frequency they had available, VMS data could not adequately describe fishing effort.

Mejias (1999) conducted a small pilot study to look at the use of VMS in a prawn fishery in the Gulf of Mexico. The study was limited to 10 boats due to problems getting fishers' support and ran for a single fishing season. In particular Mejias (1999) was interested in determining the effort during spawning. Boat activity (i.e. steaming or trawling) was determined through the use of winch sensors. These sensors gave information on the release or retrieval of the trawl nets. This sensor information was included in the VMS packet that was sent hourly.

6.3 Fine-scale effort and CPUE mapping

The majority of literature concerned with VMS is focused on the effort mapping applications. This is the simplest use of VMS data. Some studies extended the process to match the VMS data with logbook catch data to produce maps of CPUE.

In an FRDC publication, Dichmont *et al.* (2001) and Deng *et al.* (2005) reported on the use of VMS data in the Northern Prawn Fishery of Australia. One of the study's aims was to produce more accurate fishing effort distribution maps. Fishing intensity was defined as the number of times an area is swept by a trawl net. The study applied simple decision rules to define a trawl signature for the fleet (i.e. only records where, for example, trawling speed was calculated at between zero and four knots and time of day was between 8am and 6pm). Haddon *et al.* (2006) in another FRDC study again used relatively simple speed-based decision rules combined with categorical analysis to extract trawl signatures and areas of highest trawl intensity from VMS data of the Bass Strait scallop fishery. Haddon *et al.* (2006) make the important point that finer spatial resolution is possible if higher polling rates are used in VMS. This highlights the unstated difference between the objectives of compliance, the main use of VMS data, and the objectives of research/assessment uses.

Although not utilising a satellite-based VMS, the study by Marrs *et al.* (2002) raises many of the same issues as found when using VMS. Marrs *et al.* (2002) conducted a pilot study to examine the use of data loggers as tools in fisheries research. The use of data loggers serves a similar purpose to VMS, in fact the data logger could be simply thought as the equivalent of a VMS unit that is polling every 10 minutes. The study was conducted over four months in the *Nephrops* fishery in the Clyde, to the west Scotland. The data loggers were installed on eight vessels, recording the vessels position every 10 minutes. Marrs *et al.* (2002) matched the position data with logbook data and plotted landings per unit effort (LPUE) at fine spatial scale.

6.4 Resource intensity mapping

Many fisheries, including Queensland, use catch per unit effort (CPUE-) based reference points and Fishery Review Events. CPUE is often calculated by summing daily catch and relating it to effort recorded within large-scale defined grids. At this scale CPUE can be seen as a relatively poor index of abundance. Also, the targeting of high-density concentrations of the resource by fishers can lead to hyperstability (i.e. local catch rates remain high whilst the stock declines).

With the introduction of the VMS, a vessel's position can now be monitored frequently at a fine spatial scale. At this spatiotemporal scale the use of geostatistical techniques to calculate stock density may be more applicable than using average CPUE. However, many of these techniques do not perform well when trying to map a highly aggregated resource (Maravelias, Reid *et al.* 1996). These methods tend to treat some parts of highly aggregated concentrations as outliers due to the difficulty in modelling them (Rivoirard, Simmonds *et al.* 2000).

Methods such as kriging rely on linear functions (i.e. variogram models) to describe the relationship between neighbouring points, which in turn assumes a gradual change from one point to another (Rivoirard, Simmonds *et al.* 2000). This approach is acceptable if there is sufficient fishery-independent sampling coverage, which is based on a grid of systematic samples taken over a study area. For schooling species, sampling of stock density is usually carried out using sonar, echo sounders or other cost-effective remote sampling tools. However, prawns and scallops are demersal species and relatively cheap fishery-independent sampling using the above or similar tools is not feasible. Estimates of stock density or stock size are usually obtained from commercial logbook catch and effort data.

Using commercial data to conduct geostatistical analysis can be problematic. Murray (1996) provides a classic fisheries example of a highly skewed resource (Antarctic krill) where geostatistical techniques failed using commercial data. While there is position information on individual tows, the majority of fishers record daily catch. Consequently, there is little precision regarding the location of catches at a scale finer than the distance a trawler can travel in a day. In fisheries such as the Queensland deepwater EKP, for example, fishers normally trawl for approximately four hours at a time. With an average speed of around three knots, a single shot can extend for 12 nautical miles. To map catches and associated indices at spatial scales finer than a shot length, we need to estimate where along the trawl shot the majority of the catch is caught. This is possible to determine if we have a number of individual tows and associated catches that crisscross each other within a relatively short period of time and enough are contained within a relatively small region.

Often stocks are highly aggregated and disperse due to their respective biological characteristics and spatially patchy habitat (e.g. tiger prawns on the Great Barrier Reef and near-shore islands). Therefore a technique for estimating resource intensity that does not rely on linear modelling would be more useful. Probability modelling techniques such as maximum entropy analysis show particular promise.

Maximum entropy has been used for a number of fisheries applications. Vignaux *et al.* (1998) applied maximum entropy techniques to estimate fish density in the New Zealand hoki spawning fishery. The fishery is approximately 96 by 160 nm in size and the authors divided the area into 8 by 8 nm blocks for analysis. Brierley *et al.* (2003) successfully applied maximum entropy on acoustic survey data to map Antarctic krill density and biomass. Results suggested that maximum entropy is useful for situations where sample transects are evenly distributed, as compared to randomly placed transects where there may be large distances between transects. The main reason for this is that maximum entropy makes no inference outside the data range.

6.4.1 *Bayesian Maximum Entropy*

Maximum entropy analysis originated in the 18th century through the works of Bernoulli and Laplace and reintroduced initially by the work of Jaynes (1957). Maximum entropy (MaxEnt) is a technique based on probability theory for reconstructing distributions. It is particularly suited to handling noisy and sparse data in a consistent manner (Gull and Skilling 1999). The basic theory behind MaxEnt is that, lacking information about a certain quanta we assign it a probability distribution, we then choose from a range of distributions that satisfy constraints defined by our prior knowledge the distributions that maximises the entropy, S. These distributions must be both to positive and additive. For example, the fish intensity within the considered area has to be positive and the net catch of two non-overlapping cells is their sum.

The technique was first applied in the reconstruction of fuzzy images (Gull and Daniell 1978), reducing background noise whilst sharpening the major parts of an image. MaxEnt has also been used to X-ray tomography. X-rays are passed through the body from a number of directions and the absorption intensity of each X-ray line is measured. The MaxEnt method uses the information from a large number of scans to map areas of high absorption.

Whilst MaxEnt is used to assign a prior distribution to initially model fish density, Bayesian techniques are used to update the posterior distribution. Bayes's theorem in a maximum entropy context can be defined as:

$$P(hypothesis \mid data) = P(hypothesis)P(data \mid hypothesis) / P(data),$$

where $P(hypothesis)$ is the prior probability assigned to the distribution of fish density using maximum entropy, and $P(data \mid hypothesis)$ is the probability of obtaining the data given a certain prior density, also known as the experimental likelihood. In many cases this is of a Gaussian form, and $P(data)$ is the probability of the data, also known as the evidence. The posterior probability, $P(hypothesis|data)$, is the updated probability of the fish density. $P(data)$ is a normalisation constant and is used to make comparisons between possible maximum entropy solutions or hypotheses.

The main aim is to modulate the prior data to maximise the objective function (i.e. the posterior likelihood) using non-linear optimisation search algorithms.

6.4.2 *Possible limitations*

Whilst the analogy between X-ray tomography and fisheries tomography is valid there are a number of differences that need to be taken into account when applying MaxEnt to fisheries data.

Stock depletion

Tows and associated catches differ from X-rays and absorption density in that observations change the state of the image you are investigating. That is, we are removing the very thing we are measuring. This may result in a declining CPUE within an area over time, leading to an underestimate of the resource.

Factors that will affect the rate of decline would include intensity of trawling, time period of data collection and biological characteristics of the fish being trawled.

Stock movement

Time periods need to be chosen so that sufficient catch and effort data can be collected to make predictions of stock density in a relatively small area without being influenced by stock movement.

Catchability

Catchability can change in response to factors such as moon phase, time of night, prevailing weather conditions, location, and trawl intensity. Standardising the catch rate data may remove some of this variability, however a substantial amount of noise may remain in the data, lowering the accuracy of the estimate of the stock density.

Economic

Fishers usually know what catch rate is required to be economically viable. For example, if catches are historically low in some areas, fishers will not fish there unless prices are high enough to justify it. Therefore information regarding density in these areas will be lacking.

Some of these differences can be taken into account by either manipulating the data or by adding more layers of modelling into the general maximum entropy framework. Maury and Gascuel (2001) and Maury *et al.* (2001) developed an advection-diffusion-reaction model to study the effects of local overfishing on yield. There are two fish movement components of the model; a random one, a diffusion term that takes into account general dispersion; and a directional one, an advection term that follows an environmental/habitat gradient. Other possible alternatives include targeting only a high-density resource such as a spawning stock (Vignaux, Vignaux *et al.* 1998) where there is little stock movement and high trawl intensity over a short time period.

6.5 Other uses of fine-scale data

Dichmont *et al.* (2001), Deng *et al.* (2005), and Haywood *et al.* (2005) reported on the use of VMS data in the Northern Prawn Fishery of Australia. These authors also examined the use of VMS data to conduct a depletion analysis on some of the grids, comparing random and aggregated fishing effort.

Similarly Gedamke *et al.* (2004) used VMS data to 'disentangle' the effects of non-random fishing patterns in order to apply a depletion model to estimate scallop dredge efficiency (cf catchability). Bertrand *et al.* (2005) explored use of a fractal analysis of VMS-derived fleet movement as a possible tool for real time monitoring of ecosystems. Their applications described the classic 'predatory' behaviour of Peruvian purse-seiners fishing schools of anchovy.

Using a combination of VMS and logbook data, Murawski *et al.* (2005) evaluated effort and catch distribution adjacent to temperate Marine Park Authorities to show 'spill over' of groundfish from the protected area and a consequent increase in fisher revenue within 4 km of the boundary. The authors noted that such analysis was only possible because of the high-resolution vessel position data available from VMS.

Heywood *et al.* (2006) suggested the use of maximum entropy estimates of abundance within small-scale management units in a study of Antarctic Krill. The MaxEnt formalisation allows objective choice of parameters, an intrinsic calculation of errors, and enhances the conservation and management potential of sparse (acoustic) survey data.

6.6 Software

There was not much dedicated software available for analysing VMS data in fisheries. Some electronic logbook packages offer extra features such as effort and CPUE mapping but these are rather rudimentary and do not offer any solutions specific to handling VMS data (i.e. trawl signatures).

6.6.1 *OceanFARM*

This section – a review of the OceanFARM software package – begins with a brief summary of the OceanFARM's main features and options. Then we critically evaluate the package for use in this project, describing its advantages and disadvantages.

Note: *This review relates to OceanFARM's specific use in this project and is not meant as a comment on OceanFARM's general utility in other applications. Our aim is to examine the current OceanFARM version's ability to meet the specific aims of this project.*

Description

See 2.2 TerraSystems software: OceanFARM and TerraVision.

OceanFARM serves two main functions: effort-density mapping and resource-intensity mapping. These two processes are interrelated since both depend on the same analysis of the data (i.e. trawl signatures) to successfully extract line segments corresponding to trawling (trawl segments) from the data. Effort mapping simply summarises the density of the trawl segments, using the catch data as a guide to which species is being caught. For example, an effort-density map of the scallop fishery would use the trawl segments corresponding to scallop catches only, with no use of the actual amount caught. Alternatively, resource-intensity mapping uses the actual reported catch weights as well as the trawl segments to estimate the intensity distribution of the resource via the maximum entropy algorithm.

The OceanFARM package is made up of a library of functions that can be utilised in user-defined FARM macros (Visual Basic script). This provides a degree of flexibility and control for the user to define trawl signatures and allow manipulation of the data. Basically OceanFARM interfaces with tables in an Oracle database corresponding to the VMS data and the logbook catch data. OceanFARM produces a FARM data or '.FAD' file by matching logbook recorded catch with VMS position records.

Using suitable FARM macros the user may then manipulate the FAD files, or produce models for the fishery. These models can in turn be imported into MAPX® to produce effort-density or resource-intensity maps.

Advantages

OceanFARM has the following advantages for this project:

- The VMS compliance system in Queensland uses TerraVision so OceanFARM interfaces directly with the existing database provided to the project by QDPI&F.

- TerraVision has already invested time and development into OceanFARM, for example, the implementation of the MaxEnt algorithm and database interface.

- OceanFARM allows the user to define trawl signatures by writing a Visual Basic script routine that assigns a likelihood of trawling to each trawl segment.

- OceanFARM is not tied to any particular logbook system. For example, it will currently use data from ECERS, and the manually sourced CFISH logbook system. It can also be readily adapted to AFMA's electronic logbook system.

Disadvantages

The application of this technology to real VMS/catch data is still at a developmental stage. This requires the flexibility to examine various methodologies and modify the approaches OceanFARM uses. In its current form, OceanFARM does offer some flexibility and customisation via the use of the FARM macros, but many of the procedures are fixed and inaccessible to us as end users.

The main aspects of OceanFARM where we required greater flexibility are as follows:

a) Map output

OceanFARM creates effort and resource maps in MAPX. For use in this project we would prefer to produce maps in ArcMap. This is, firstly, because quality professional maps are an important tool in this project and ArcMap provides a superior output quality.

Secondly, one of the aims of this project is to provide information and maps to the QDPI&F. Fisheries managers in QDPI&F predominately use ArcMap (as it is supported by the Department) and is therefore the most convenient format.

Thirdly, ArcMap provides a range of tools to further analyse the output such as the ability to spatially query the output and overlay other information.

OceanFARM does not provide a way to extract the model information to plot in a third party application such as ArcMap. There is the option within MAPX to view the map data, but this feature has not been implemented yet and the values displayed are incorrect.

b) Grid size

Currently the number of grids that OceanFARM can model is 100x100. For small-scale local maps this is quite suitable. However, one of the project outcomes was to produce full Queensland maps at the one-minute resolution, which corresponds to a grid size of 1140x840. One solution would be to manually produce a patchwork of smaller 100x100 maps and join them together to produce a full map of the coast.

c) Definition of catch

When matching the logbook catches to the VMS position data we require some flexibility on the rule for joining the two databases. For example, we may want to match catch for targeted species only, to define as species the records that are a certain percentage of total catch for that day. This could be accomplished by manipulating the actual catch records so that OceanFARM only considers the catch record we want to use. However, this manipulation may not be convenient/easy for large databases and would require several versions of the logbook database table to allow different rules to be applied.

d) Anchorages

A major flaw was discovered in our early maps relating to when a boat enters or leaves an anchorage. The resulting calculated speed is biased downward giving the indication of trawling when in fact the boat is steaming. For example, consider a boat steaming at seven knots toward a port and is polled by VMS. Let us say that 15 minutes later the boat enters port and spends the remaining 45 minutes, until the next VMS poll, stationary. The calculated speed for this boat would be then 1.7 knots – and so deemed to be trawling. This problem is exacerbated since the paths taken to port/anchorages are often used repeatedly by many boats. The problem manifests itself by extreme effort hotspots on the entry and exit to ports or anchorages, as seen in the corresponding resource map (Figure 6.1).

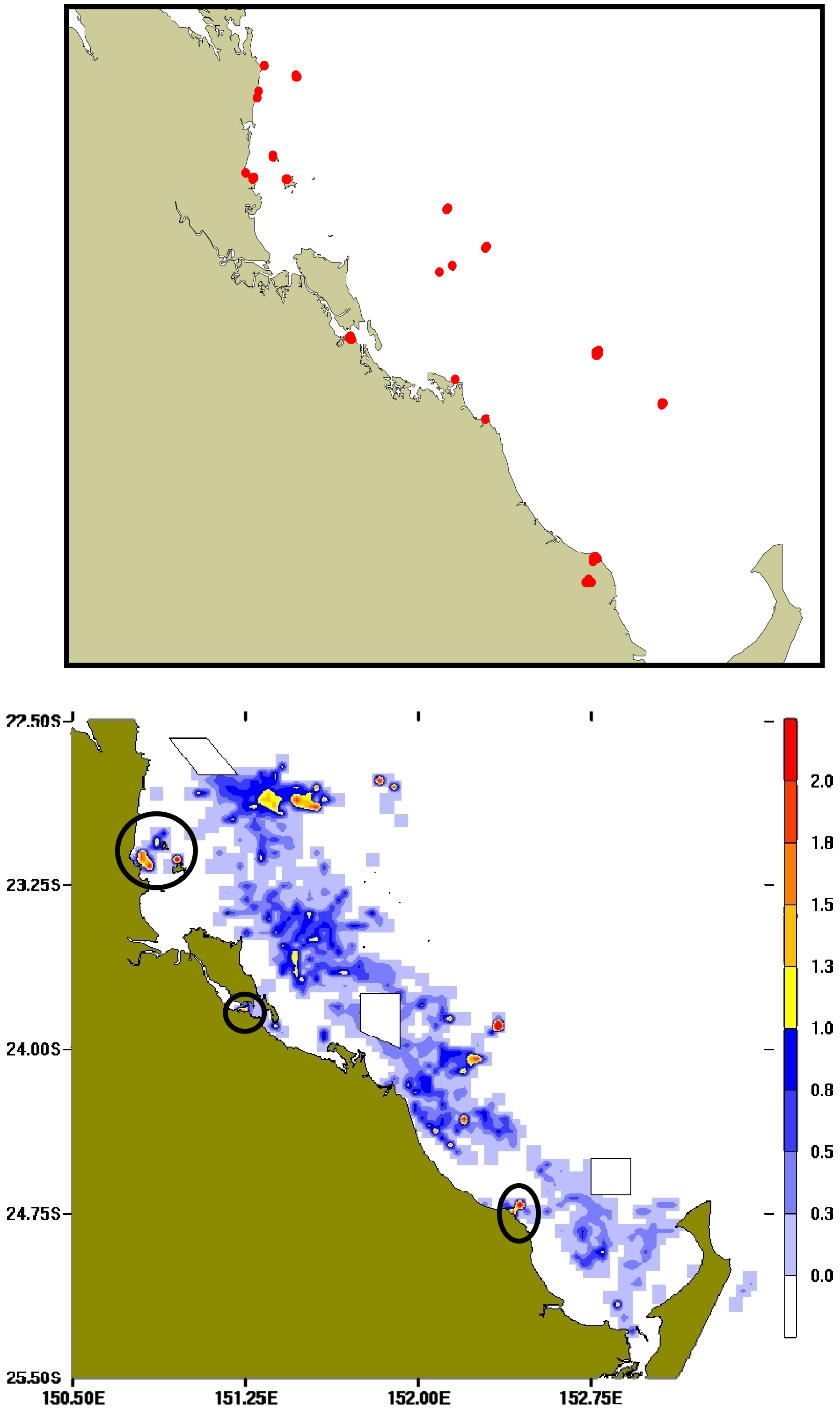


Figure 6.1 A comparison with known anchorages (top) and an example of a resource intensity map from OceanFARM for the Scallop fishery (bottom) with possible anchorage errors highlighted.

e) Maximum entropy

Some technical issues must be overcome when applying maximum entropy to fisheries data. Possible solutions are well covered in Chapter 9. However, in the version supplied to this project, OceanFARM does not allow modification of the MaxEnt algorithm.

6.6.2 *GEOCRUST 1.0*

Afonso-Dias *et al.* (2002) have produced a stand-alone GIS package called GEOCRUST 1.0 to store, analyse, and display VMS and landing data of the South-Southwest Portuguese crustacean fishery. GEOCRUST is made up of seven modules:

- Map and edit original VMS data
- Analyse the speed of each vessel
- Define the boundaries of all fishing trips by a vessel in a year
- Identify and define trawl tows within a fishing trip
- Produce maps of fishing effort of LPUE
- Exploratory data analysis and export of data
- Recreate activity of a boat or group of boats.

The package seems quite useful for initial simple analysis and exploration of the data. The study by Afonso-Dias *et al.* (2002) was quite small compared to the number of vessels in this project. In addition, the software does not provide resource mapping functionality required by this project, and therefore it was not considered.

7 MAPPING THE SPATIAL INTENSITY OF FISHING EFFORT USING SPEED FILTER METHOD

Norm Good, David Peel and Mai Tanimoto

7.1 Introduction

Mapping trawl effort precisely and accurately is a major objective in fisheries management and stock assessment (Booth 2000; Marrs, Tuck *et al.* 2002). Improving the spatial resolution of catch and effort records will lead to more reliable assessments of stock abundance and resource distribution. This is especially important when using CPUE as an index of abundance for stock assessment models.

However, the assumption that conventionally derived CPUE is proportional to abundance has been disproved many times (Hilborn and Walters 1992). The targeting of aggregated populations by cooperating fishers can create hyperstability in a CPUE series, especially when viewed on a relatively large spatial scale.

Fisheries management is increasingly concerned with the impact of trawling on benthic habitat, the associated faunal communities and incidental bycatch species (Burridge, Pitcher *et al.* 2003; Koslow, Gowlett-Holmes *et al.* 2001). To objectively measure trawl impacts, fine-scale trawl effort data are required to determine trawl intensity and distribution. Relying on large-scale logbook data to measure impacts may give biased answers, especially in aggregated fisheries such as prawn and scallop.

A number of studies have investigated the fine-scale distribution of trawl effort using a range of sample tools. On-board position data loggers measure trawl speed directly or at a very fine temporal scale, and fishery-independent surveys record the position of the start and end of each trawl shot. These studies have explored trawl effort in relation to bottom disturbance (Marrs, Tuck *et al.* 2002; Rijnsdorp, Buys *et al.* 1998), fleet dynamics (Bene and Tewfik 2001; Dorn 2001; Fletcher 1992; Hampton and Fournier 2001; Hilborne and Ledbetter 1979; Maury and Gascuel 2001; Pet-Soede, Van Densen *et al.* 2001; Rijnsdorp, Dol *et al.* 2000; Rijnsdorp, van Mourik Broeman *et al.* 2000), stock assessment (Booth 2000) and identifying spawning stocks (Begg and Marteinsdottir 2003). However, the long-term ability to measure fine-scale trawl effort is hampered by the need to constantly conduct these types of surveys. In addition, as these studies mainly use a sample of the fleet, sampling error is incorporated into subsequent modelling.

Continuous trawler position information has been recently collected in some Australian fisheries. In the Northern Prawn Fishery, Dichmont *et al*. (2001) mapped fishing intensity by applying trawl track analysis to VMS data. The system collected position information for each trawler in the fleet at regular time intervals. Fishing intensity was defined as the number of times an area is swept by a trawl net. The study applied simple decision rules to define a trawl signature for the fleet (i.e. only records where trawling speed was calculated between zero and four knots and time of day was between 8 am and 6 pm). However, no effort was made to further refine the decision rules to obtain a more accurate signature of trawling. Larcombe *et al.* (2001) studied trawl intensity in the South-East Fishery to assess marine disturbance. They used start and finish positions of a trawl shot from logbook data to define a trawl track. No formal definition of a trawl signature was made as it was assumed that trawling only took place between the start and finish locations. Thirty per cent of the data covering the period 1995–99 contained 'spurious' tracks and was excluded from analysis. A number of trawl effort maps were produced, and whilst they mapped trawl effort accurately, the spatial precision remained dependent upon accurate reporting by fishers.

Stock assessment and effort mapping of Queensland trawl fisheries relies primarily on commercial catch and effort logbook records (e.g. Dichmont, Haddon *et al.* 1999; Williams 2002). The spatial reporting of these records varies considerably but is currently reported to an approximate 6 x 6 nm^2 grid. Established in December 2001, the VMS has continuously collected hourly position information ever since for about 480 of Queensland's otter trawlers.

In the current study, a number of decision rules and techniques were developed to determine when a vessel is trawling (a 'trawl signature') from the VMS data. The chronology of the trawl signature development is summarised as follows:

Phase 1: determine trawl activities based on a single trawl speed rule

- fixed cut-off speed of 2 m/s (3.8 knots) for all fisheries

Phase 2: determine trawl activities based on a fishery-based trawl speed

- fixed cut-off speed for each fishery estimated from the mixture model

Phase 3: determine trawl activities based on Hidden Markov Model

- A probabilistic allocation of trawling vs steaming using the Hidden Markov Model (HMM).

Initially, a single signature of trawl speed for the whole fleet was developed (Phase 1). Secondly, separate signatures were developed for four otter trawl sectors – the tiger/endeavour prawns, the EKP, the scallop and the banana prawn fisheries – to account for likely differences in trawling behaviour (Phase 2). Lastly, probabilistic allocations of trawling activities were considered rather than an approach based on some form of outright classification. Decision rules for vessels at anchor, approaching or departing anchoring points, and other activities not associated with trawling were also developed along with these phases. This chapter explores filter methods (cut-off rules), and Phase 3 is discussed in the next chapter. Trawl effort maps for these three sectors (tiger/endeavour, EKP and scallop) are presented at the fine spatial scale of 1 x 1 nm^2 grid.

7.2 Methods

7.2.1 *Areas*

Effort maps were created for the EKP, the scallop, the tiger/endeavour prawns fisheries and total trawl effort in selected fisheries for Queensland. Note that the maps for the banana prawn fishery were not considered as it is a relatively minor fishery in Queensland. For this study the spatial extent of each fishery was delimited to include only statistical grids to which catch has been reported. Figure 7.1 shows the historical extent of the majority of the otter trawl fisheries based on commercial 30 x 30 nm^2 logbook data.

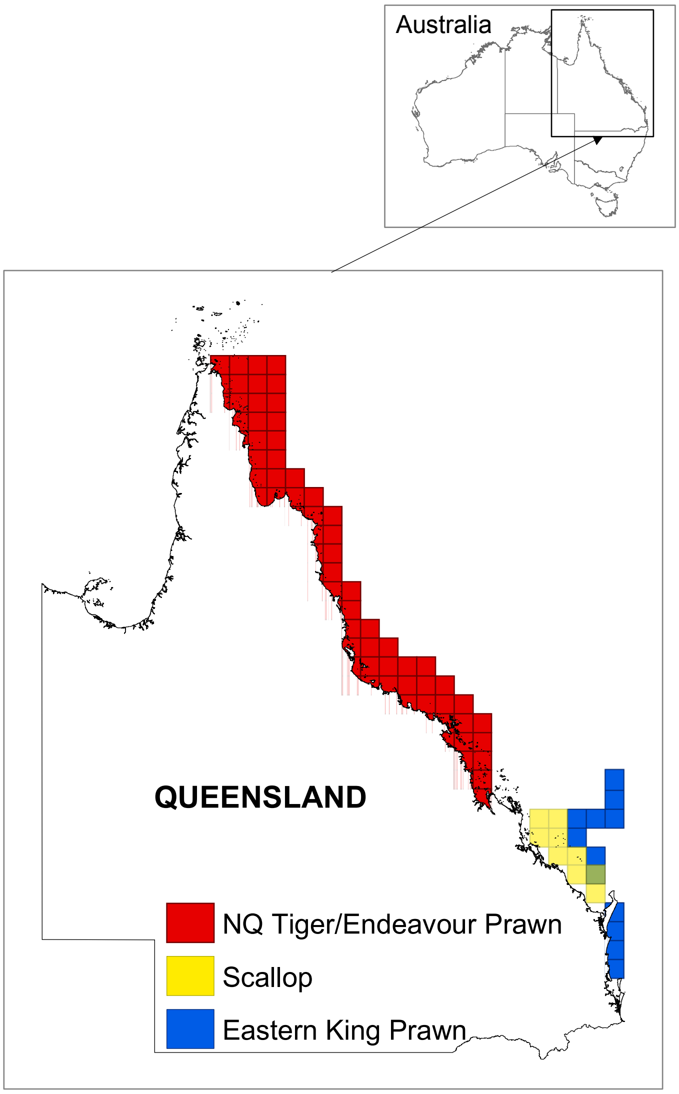


Figure 7.1 Spatial distribution of catches for each otter trawl fishery in Queensland. CFISH grids representing extent of respective prawn and scallop stocks. Lightly shaded grid represents overlap of EKP and scallop stock.

7.2.2 *Data*

VMS data

The position of a vessel is transmitted hourly to an INMARSAT satellite from a ship-based transponder. This information is sent to a land-based station and then to the VMS unit of the QDPI&F. VMS data were originally collected voluntarily by a number of boats from December 1999, with a substantial number of polls – about 20,000 – being recorded from April 2000 until December 2000, when VMS became compulsory. The temporal trends in poll count are indicative of seasonal activities in the fisheries (Figure 7.2).

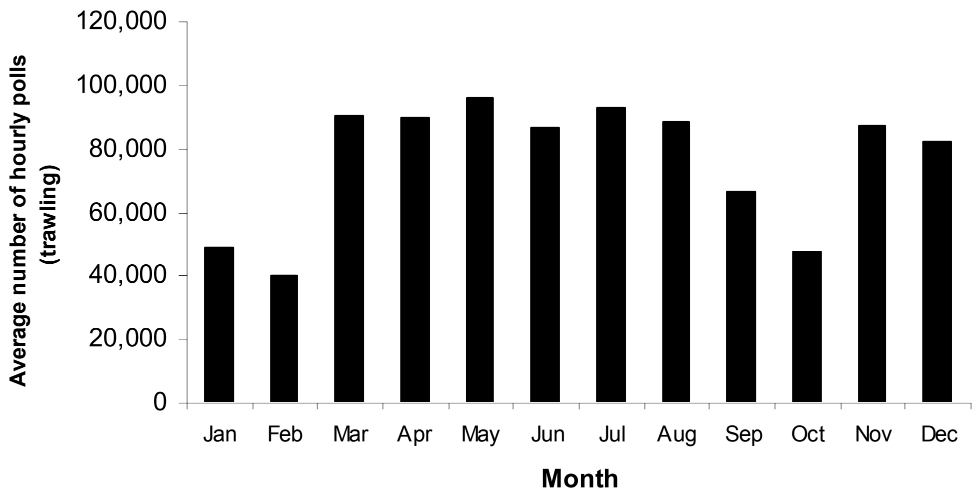


Figure 7.2 Monthly average number of hourly VMS trawling polls (Dec 2000–Mar 2003)

GPS data

As part of the project, GPS track data were obtained from a number of fishers over a range of fisheries. These data were used to test the methodology developed to extract trawl information from VMS data. The polling frequencies of the GPS data are such that the data can be considered as a substitute for the true path of the vessel and hence provide more indicative estimates of the true intensity of trawl effort of the vessel. However, since we have no actual indication of actual behaviour (e.g., trawling versus steaming) there will still be some error in the effort intensity estimated from GPS data. This error is acceptable as our aim was simply to provide a benchmark against which to test our ‘simulated’ polled VMS data.

To better represent the full spectrum of fishing behaviour we required good spatial coverage of GPS data including reasonable samples from each major fishery and a reasonable number of fishers. Close examination of the collected data showed a large proportion did not include time information so we were unable to use the data directly in the simulation study. To overcome this we attempted in many of the cases to estimate the polling frequency based on the average distance travelled at each poll and an estimate of typical trawling speed. Obviously this is not ideal but as the purpose is simply to produce data that reasonably represent trawling behaviour it was considered to be appropriate. In some cases the fisher's GPS units are set to take polls based on a set distance travelled rather than a timed poll. In these cases we could not accurately estimate polling frequencies, therefore these data were removed.

High-frequency VMS

Another source of accurate high-frequency polled data is the VMS data themselves. The database contains a number of instances where units have polled as frequently as once per minute. These data are collected irrespective of trawler activity, unlike the GPS data that are generally collected only during trawling. Therefore the high-frequency VMS data, as well as augmenting the spatial coverage of the GPS data, also provide a good representation of steaming as well as trawling. Figure 7.3 shows the coverage obtained from the combined data collection.

To get greater spatial/boat coverage and more steaming behaviour we included all VMS data of a polling frequency greater than 15 minutes. This increased the amount of data, but could bias the results slightly as the resulting data included a disproportionate number of scallop vessels due to increased polling frequency near scallop closures.

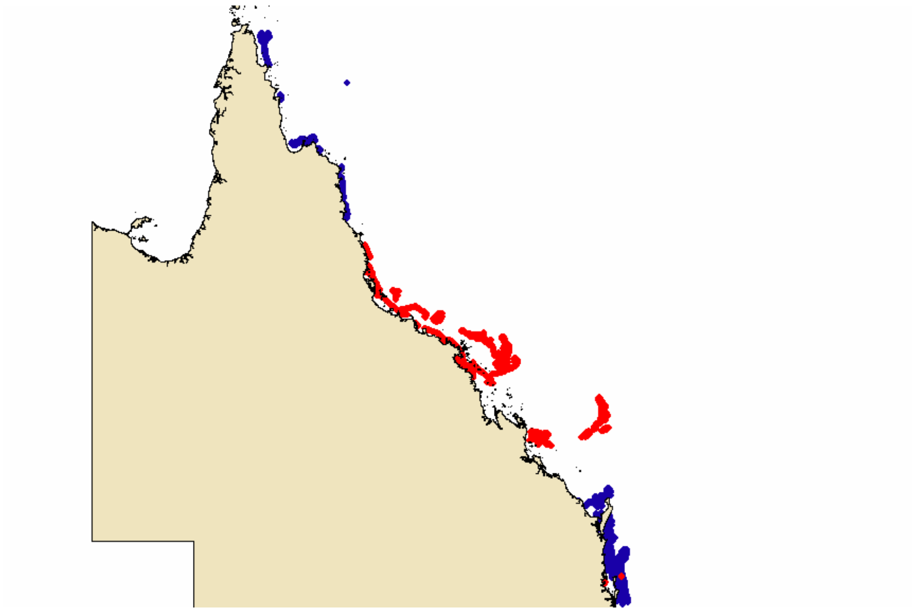


Figure 7.3 GPS (Blue) and high-frequency VMS (Red) data coverage

Logbook effort data

Logbook catch and effort data from the commercial CFISH trawl database were extracted by species. For fishery-specific trawl signatures, catches of EKPs (*Penaeus plebejus*), tiger/endeavour prawns (*Penaeus esculentus*, *P. semisulcatus*, and *P. monodon*/*Metapenaeus endeavouri* and *M. ensis*), banana prawns (*Fenneropenaeus merguiensis*) and scallops (*Amusium japonicum* and *A. pleuronectes*) were matched to VMS records by unique vessel number. To extract all trawl effort, catches of all trawl species were matched to VMS records.

7.2.3 *Decision rules for defining a trawl signature*

Trawl signatures which identify when a vessel is trawling can be derived by applying a set of decision rules to VMS data. These rules may include:

- Time-based rule
- Speed-based rule(s), and
- Location-based rule.

In this section, actual VMS data were used to graphically demonstrate the basic process to produce the effort maps using these rules. Figure 7.4 shows the VMS data for a single vessel over the period of a year. Note that these data have been transformed and rotated, and the artificial land structure has been added to ensure confidentiality (see Appendix C for Confidentiality Deed).

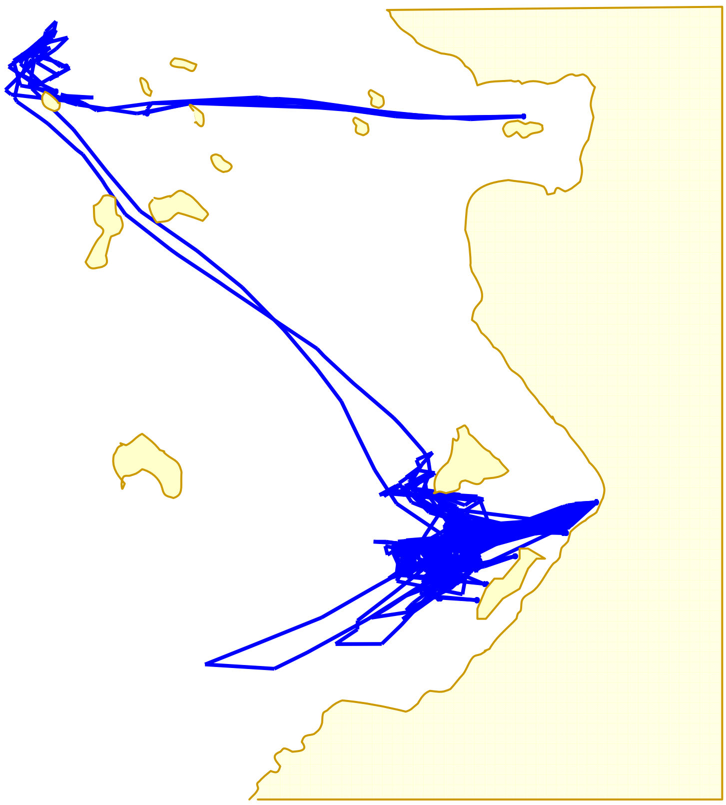


Figure 7.4 Plot of sample VMS data for a single vessel.

Time-based rule: Remove polls from non-trawling times

The time-based rule removes all polls corresponding to non-trawling times (e.g. times when the boat will most definitely not be trawling). For night-time only fisheries such as scallop and most prawns, the daylight hours are classed as non-trawling time, i.e. only polls between 5 pm and 7 am were selected as trawling times. However, these times change for particular fisheries and time of year. As many fishers work according to a dusk-to-dawn cycle an astronomical algorithm was incorporated into queries for selecting polls. All polls over a 24-hour period were selected for the banana prawn and stout whiting (*Silago robusta*) fisheries, as both fisheries can be day and night operations. The effect of this step is shown in Figure 7.5.

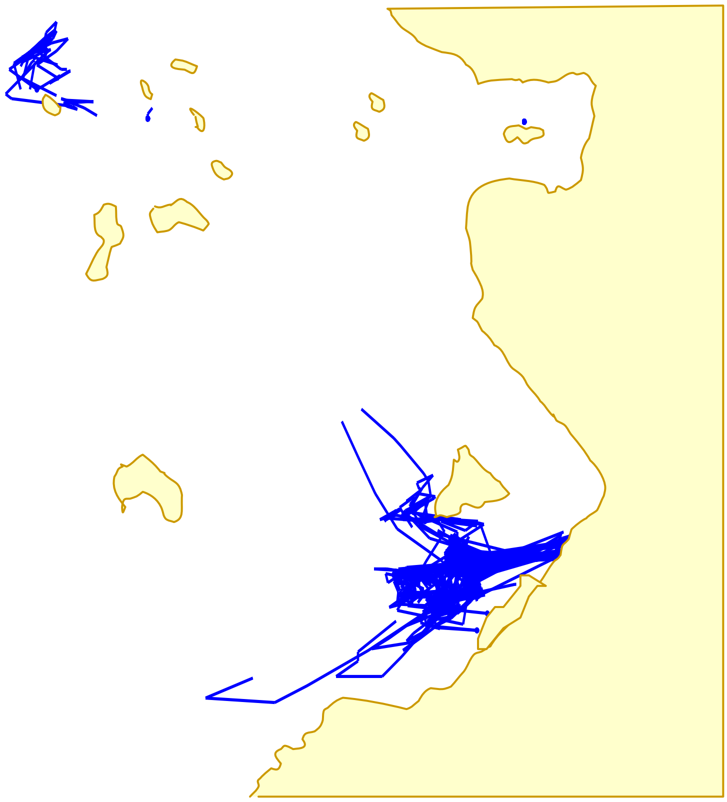


Figure 7.5 Plot of data after polls from non-trawling times have been removed.

Speed-based rule 1: Remove polls corresponding to steaming

The most crucial variable to determine trawling activity is the calculated speed. One way of determining vessel activity by probability is to fit a two-component mixture model of univariate normal distributions to the calculated vessel speed. A mixture model is a statistical model of heterogeneous data particularly useful in cluster analysis. In this application a mixture model can be considered as probabilistically clustering the historical speed data into two groups – trawling and steaming. The model can be fitted to the historical speed data using an unsupervised learning 'algorithm' and then new data can be allocated based on this model as in a discriminant analysis context.

To describe the mixture model in more detail, let $x_1,\ldots x_n$ denote the calculated speed data resulting from n polls, then the data are assumed to be distributed as:

$$f(x;\theta) = \pi_1 N(x\ ;\ \mu_1, \sigma_1) + (1-\pi_1) N(x\ ;\ \mu_2, \sigma_2)$$

where μ_1 and μ_2 correspond to the mean calculated speeds of trawling and steaming respectively. Similarly, σ_1 and σ_2 correspond to the respective variances. The parameter π_1 is the mixing proportion.

The mixture model is fitted via the Expectation-Maximisation (EM) algorithm (Dempster, Laird *et al.* 1977), providing posterior probabilities of trawling/steaming. Basically, this involves two circular steps: the Expectation (E-step) and Maximimisation (M-step) steps. Letting g denote the number of components then the $(k+1)^{\text{th}}$ step of the EM algorithm for this application is given as follows:

E-Step

The posterior probability (i.e. the probability for each point being in each group) is estimated, for $i = 1,\ldots n; j = 1, \ldots g$

$$w_{ij}^{(k+1)} = \frac{\pi_j^{(k)} N\left(x_i\ ;\ \hat{\mu}_j^{(k)}, \hat{\sigma}_j^{(k)}\right)}{\sum_{m=1}^{g} \pi_m^{(k)} N\left(x_i\ ;\ \hat{\mu}_m^{(k)}, \hat{\sigma}_m^{(k)}\right)}$$

M-Step

The unknown parameters are estimated, for $j = 1,\ldots g$

$$\hat{\pi}_j^{(k+1)} = \frac{\sum_{i=1}^{n} w_{ij}^{(k+1)}}{n}$$

$$\hat{\mu}_j^{(k+1)} = \frac{\sum_{i=1}^{n} w_{ij}^{(k+1)} x_i}{\sum_{i=1}^{n} w_{ij}^{(k+1)}}$$

$$\hat{\sigma}_j^{(k+1)} = \frac{\sum_{i=1}^{n} w_{ij}^{(k+1)} \left(x_i - \hat{\mu}_j^{(k+1)}\right)}{\sum_{i=1}^{n} w_{ij}^{(k+1)}}$$

The steps are repeated until some form of convergence occurs based on the change in the log-likelihood.

The relationship between calculated vessel speed and fishing activity may be affected by trawl track curvature (e.g. straight lines versus irregular wiggles), length and frequency of tows (e.g. a single long tow versus numerous short tows). To account for these differences in boat behaviour, separate trawl signatures are required for each fishery. In addition to the four Queensland fisheries in the study (i.e. scallop, EKP, tiger/endeavour prawns, and banana prawn), we also included a ‘default’ fishery. Trawl signatures were assigned to the default fishery when a tow did not obviously belong to any of the four main fisheries.

Using a mixture model approach provides an automated method to determine a speed-based trawl signature hence a large number of models can be fitted. This allows us to take the natural progression from fishery-based models to individual vessels. Vessels differ across the fleet in size and power, and individual skippers may fish differently. It is therefore advantageous to allow a separate model for each vessel in each fishery.

The critical point or boundary between the trawling and steaming groups is taken as the upper trawl signature speed. This cut-off speed corresponds to the point where the probability that we discriminate the vessel as steaming is equal to the probability that we believe it is trawling, i.e. 0.5 probability of trawling.

Line segments were removed when the calculated trawl speed exceeded the cut-off threshold (Figure 7.6). Much of the steaming occurs during daylight hours, and has consequently already been removed. The effect shown in Figure 7.6 is therefore not as pronounced as the previous filtering method.

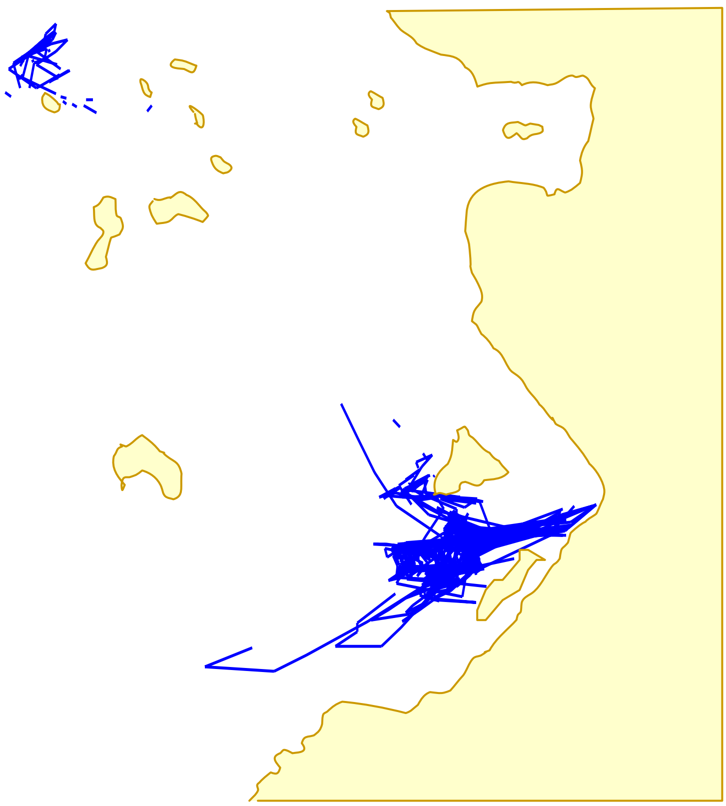


Figure 7.6 Plot of data after polls corresponding to steaming have been removed.

Speed-based rule 2: Remove polls from stationary vessels

Any line segments corresponding to stationary vessels were removed, as seen in Figure 7.7. Due to the truncation of the VMS and movement (e.g. due to drifting), even a stationary vessel may register a non-zero speed. To take this into account all vessels with a calculated speed of less than 0.02 m/s (0.04 knots) are considered as stationary. This may remove legitimate trawling polls when a vessel doubles back on its own path and by chance the hourly poll lands on top of the previous hour's poll.

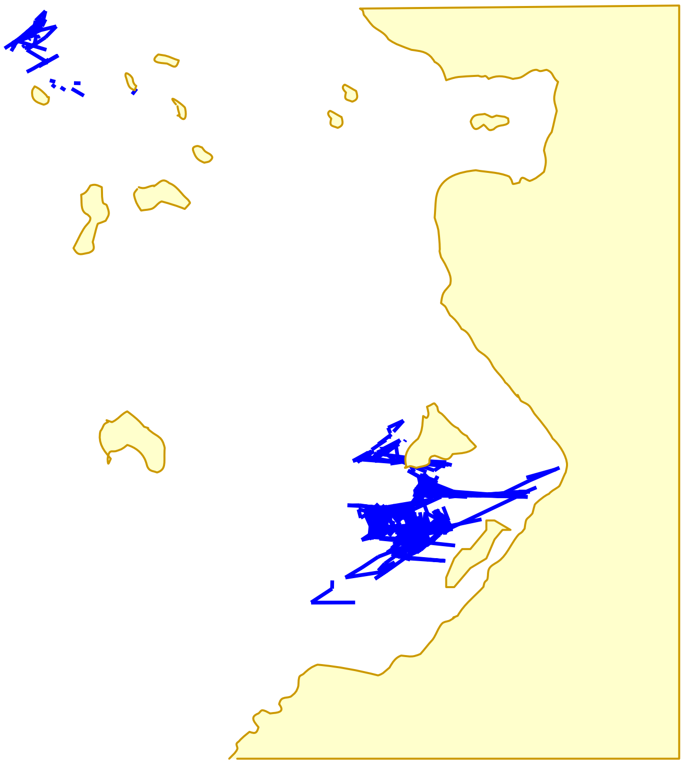


Figure 7.7 Plot of data after polls from stationary vessels have been removed

Location-based rule: Remove vessels entering/exiting ports and anchorages

Initial maps made with speed- and time-based rules applied were validated with fishers and other fellow researchers. It was quite apparent that there was an over representation of effort in ports and anchorages. In these cases a boat may have been steaming for part of an hour while either entering or leaving port. This can produce a calculated speed within the trawl speed range. To overcome this problem, a decision rule to identify polls entering/exiting port or anchorage was applied. Basically, polls immediately before or after a period of inactivity were considered transition polls and therefore removed.

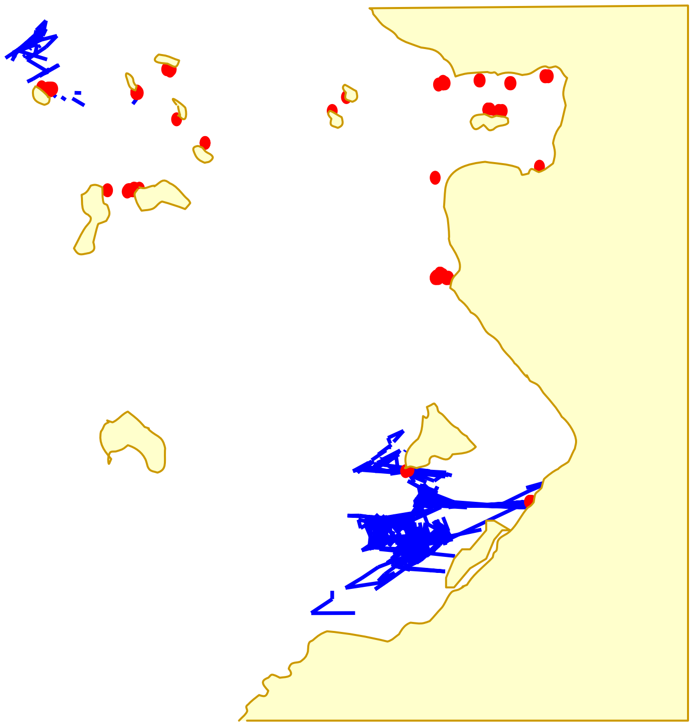


Figure 7.8 Plot of anchorages defined from VMS data

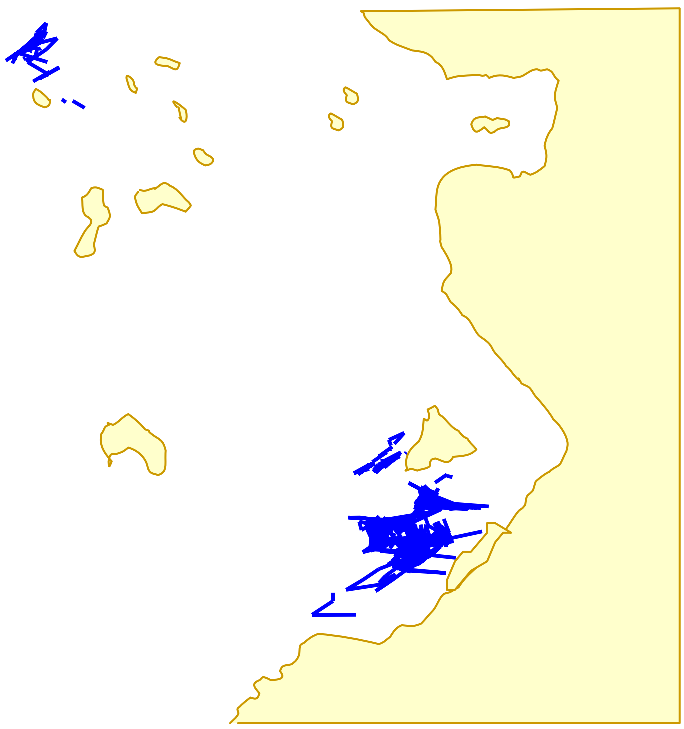


Figure 7.9 Plot of data after vessels entering/leaving anchorage areas removed

7.2.4 *Effort maps*

After the application of the decision rules a series of points mostly representing hourly trawls were obtained. To improve the fine-scale resolution a line was drawn between two polls to roughly approximate a trawl track, and effort was apportioned along the resulting line segment. Effort was then calculated as the proportion of a line segment contained within a grid (see Figure 7.10 for the illustration). The grid on the top right originally contained all one hours effort. Subdividing the line segment into 60 allocates effort into four grids instead of one. Figure 7.11 shows the resulting effort-density map when this procedure is applied to the example boat.

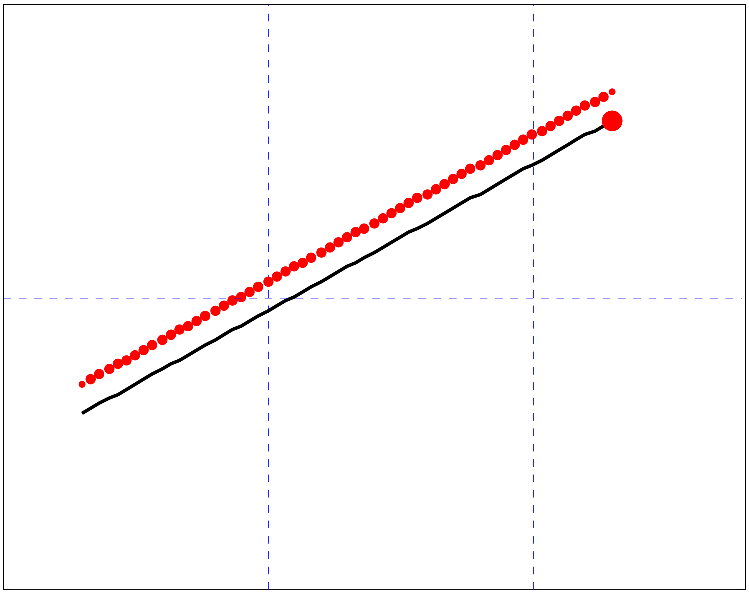


Figure 7.10 Example of proportional allocation of effort to grids by dividing a one-hour line segment into 60 one-minute segments.

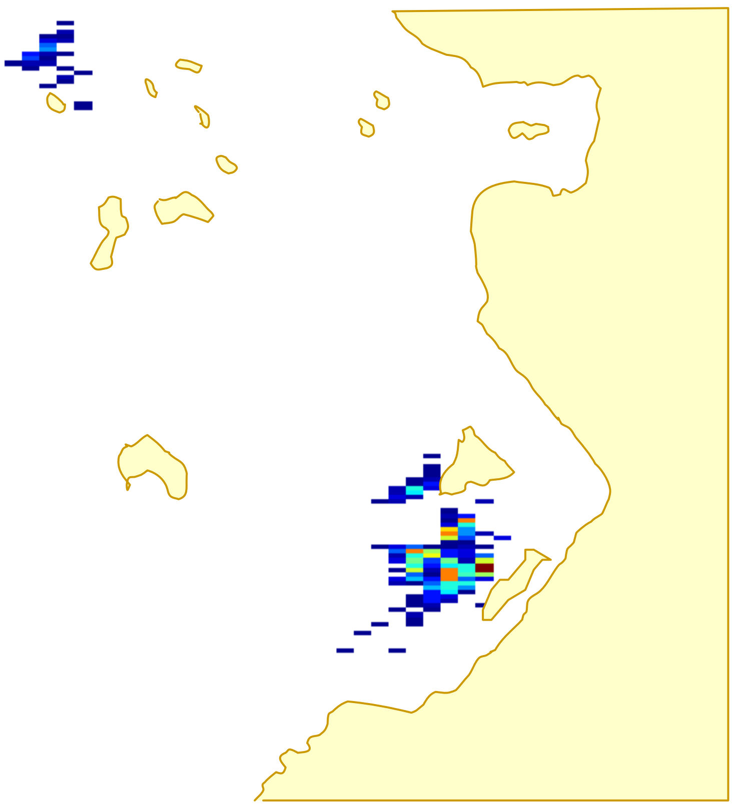


Figure 7.11 The resulting effort density map

7.2.5 *Catch maps*

To allocate catch to trawl tracks, daily logbook records were assigned to line segments in the same way as effort was allocated. That is, catch was distributed in proportion to the length of the line segment contained within a grid.

7.2.6 *Simulation study: validating, testing of speed-based rules*

Two measures were used to compare the performance of various trawl signatures. Firstly, the mean squared error (MSE) of the calculated speed between the grid trawl intensity derived from true (high-frequency polled) and data artificially polled at one hour was estimated on an hourly tow-by-tow basis. Secondly, we considered the problem in a classification context i.e. a trawl signature is simply a rule to classify the hour tow as trawling or steaming. We could then estimate the percentage of the misclassification error for a given hour-long tow. For example, if the vessel was considered to be trawling (based on the speed-based rules) and the high-frequency data showed that the vessel was at steaming speed for 15 minutes during that time, then the misclassification error would be 25%.

By repeating this process for different random hourly polls (i.e. bootstrapping) we can establish variances on these error measurements. It should be noted that the errors are not direct estimates of the true error since we are using 15-minute polling to represent true behaviour and what we identify as 'true' trawling using our artificial definition may not actually be trawling. However, the 15-minute polls will generally be more accurate than hourly polls therefore the errors provide a relative measure to compare our various trawl signatures.

In the following sections we examine the effect and validity of the upper cut-off speed in our trawl signatures/rules. Only the VMS fine-scale data were used in this section as a large proportion of the GPS data did not have species/catch information associated with it and fishery-based rules could not be applied.

7.3 Results

7.3.1 *Trawl signature*

Speed-based rule

A histogram of calculated trawl speed from each fishery is shown from Figure 7.12 to Figure 7.14. A number of distinct patterns can be found by examining these figures. Firstly, a hump or mode at about 1.5 m/s corresponds to trawling activity, and secondly, a hump at approximately 3 to 4 m/s corresponds to steaming vessels. Note that the time-based rule was first imposed on the data. The data used were therefore for the trawling period only (e.g. scallop data were for night only and banana prawn were from during the day). This explains why the hump corresponding to steaming is much lower than the trawling mode. Secondly, it is notable that the histograms for the tiger/endeavour prawns and banana prawn fisheries have an extra mode or peak as the calculated speed approaches zero. It should be noted that these values do not correspond to actual trawl and steaming speeds as they are calculated speed (i.e. the distance travelled since the last VMS poll is used). This will generally give lower speeds than the actual speed.

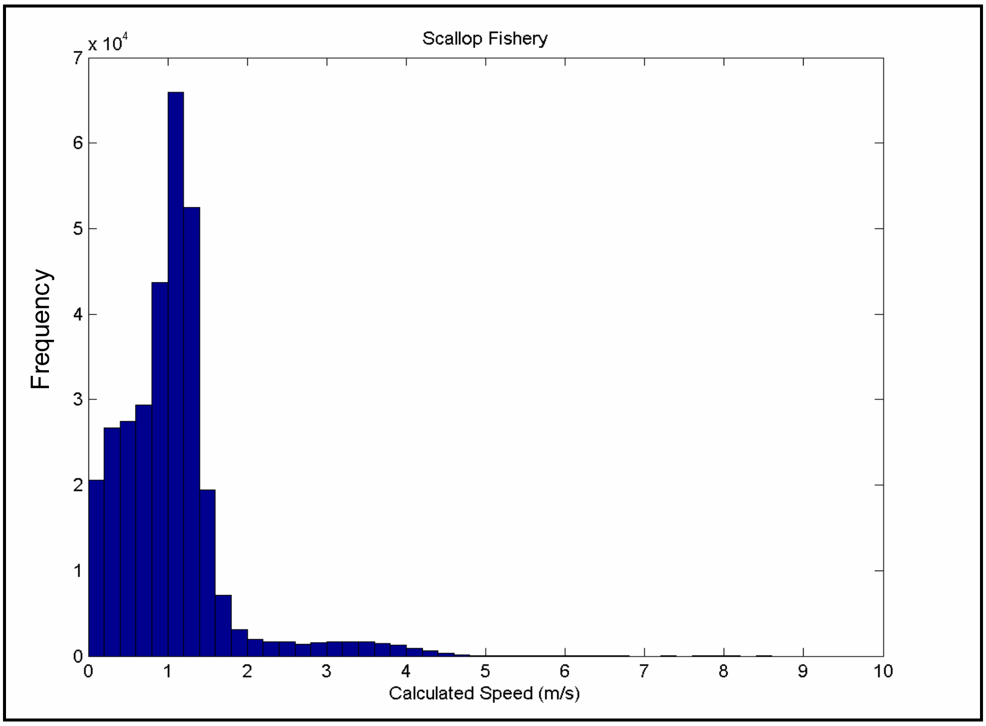


Figure 7.12 Histograms of calculated speed for the scallop fishery 2000–2003

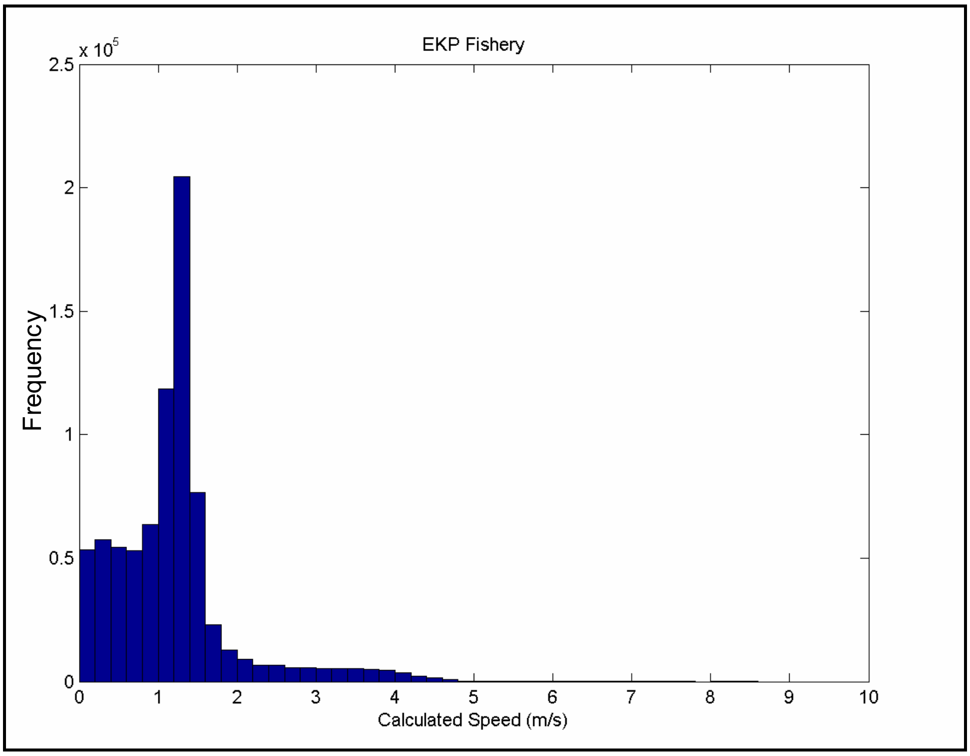


Figure 7.13 Histograms of calculated speed for the EKP fishery 2000–2003

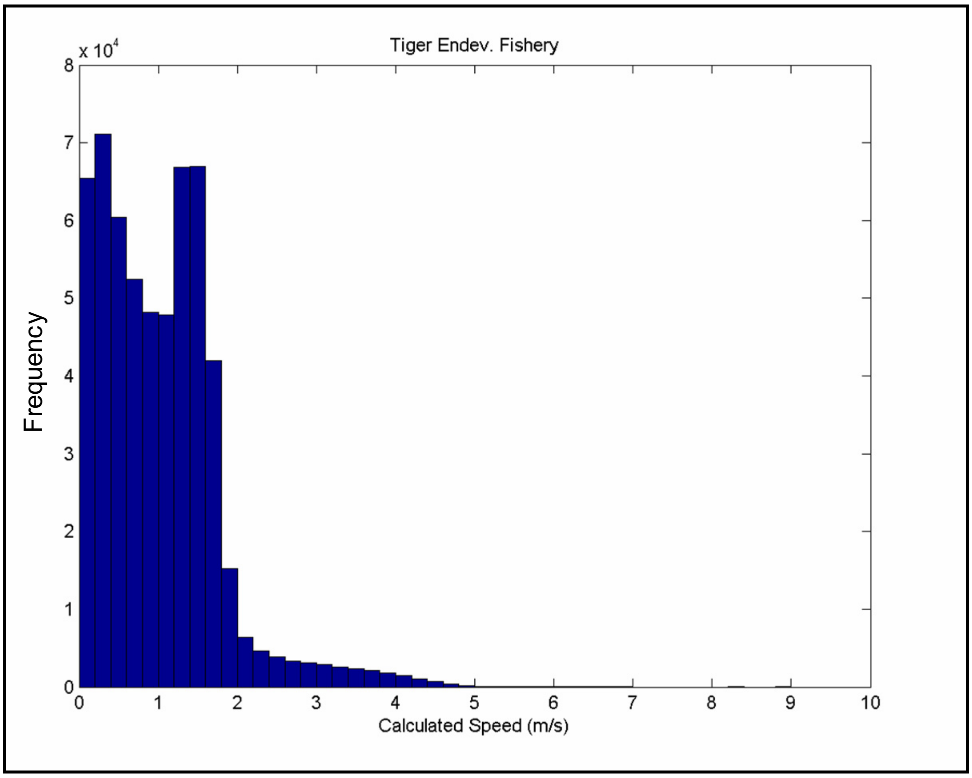


Figure 7.14 Histograms of calculated speed for the tiger/endeavour prawns fishery 2000–2003

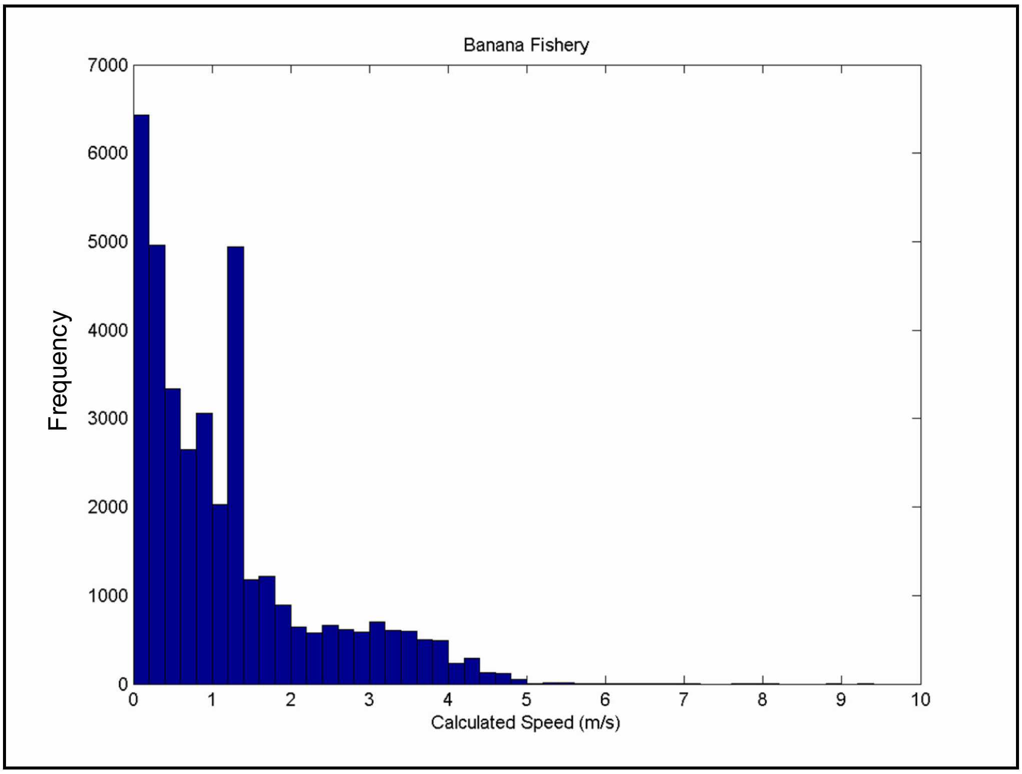


Figure 7.15 Histograms of calculated speed for the banana prawn fishery 2000–2003

A mixture model has been fitted to each fishery and the results are shown in Table 7.1. The banana prawn fishery tends to have the lowest cut-off speed and the EKP the highest.

Table 7.1 Results of estimated upper cut-off speed for each fishing sector by fitting various mixture models to respective historical calculated speeds.

Fishery	2-component normal
Scallop	2.21 m/s
EKP	2.37 m/s
Tiger/Endeavour Prawns	2.26 m/s
Banana Prawn	1.68 m/s
Default	2.41 m/s

Error estimates for various upper cut-off speeds for each fishing sector are shown from Figure 7.16 to Figure 7.19. A distinct minimum can be seen and both error measures are consistent for the scallop fishery (see Figure 7.16). In this case the optimum cut-off seems to be approximately 2.1–2.4 m/s, while the MSE seems to indicate 2.1 m/s. The upper cut-off speed of 2.21 m/s estimated from 2-component mixture model is in the acceptable range.

The minimum cut-off speeds for the EKP fishery (see Figure 7.17) were not as distinct as for the scallop fishery. If both error measures are considered, a cut-off speed approximately between 2.0 and 2.4 m/s would be appropriate. This supports the estimated cut-off speed from the 2-component mixture model (2.37 m/s).

For the tiger/endeavour prawns fishery it would seem that a cut-off of 2.0 m/s would minimise error (see Figure 7.18), which indicates that the cut-off speed estimated from the mixture model is possibly too high.

Figure 7.19 shows the result for the banana prawn fishery. If both error measures are considered equally, it seems an approximate cut-off of 1.6 would be appropriate.

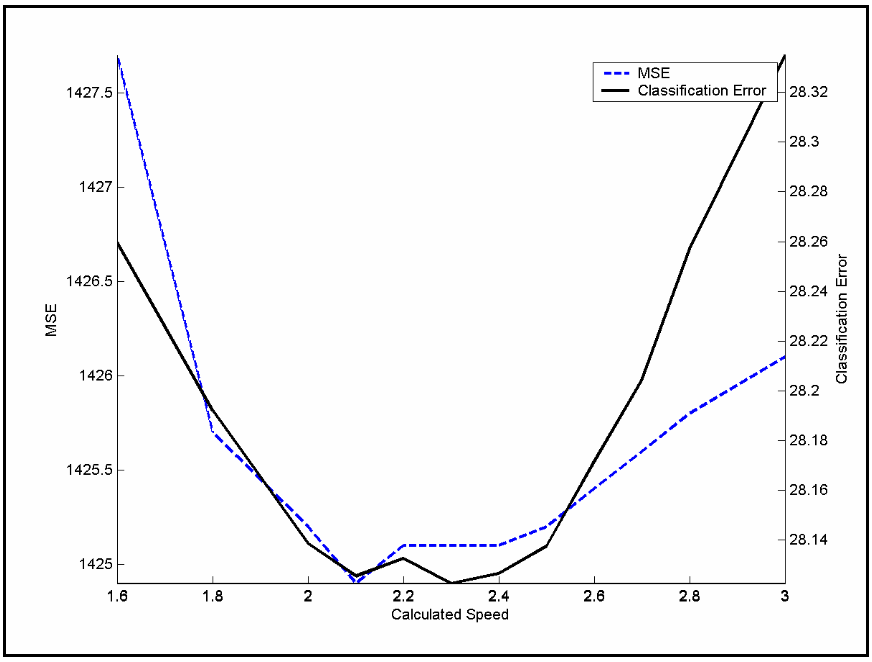


Figure 7.16 Error measures for various upper cut-off trawl speeds (m/s) in the scallop trawl signature

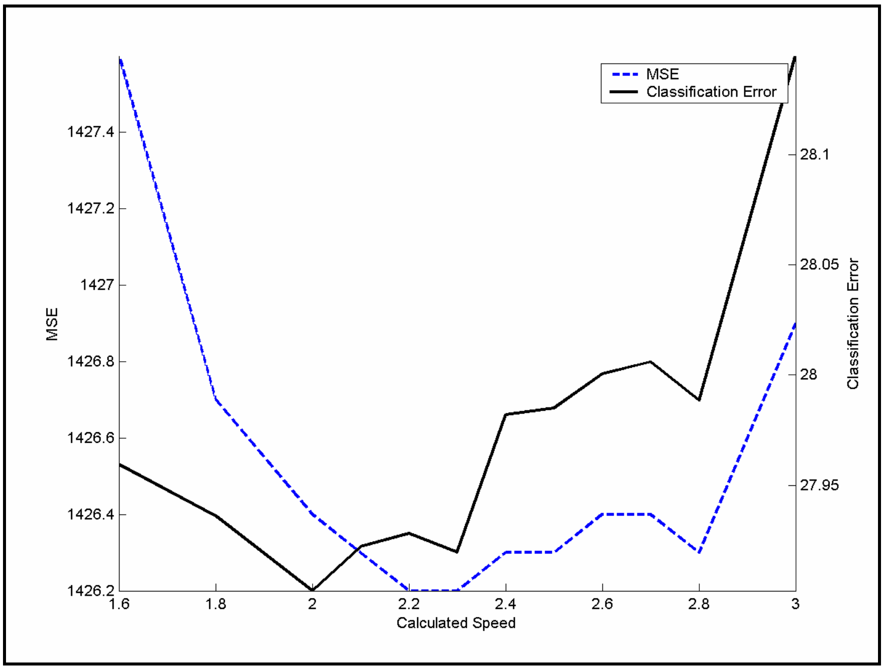


Figure 7.17 Error measures for various upper cut-off trawl speeds (m/s) in the EKP trawl signature

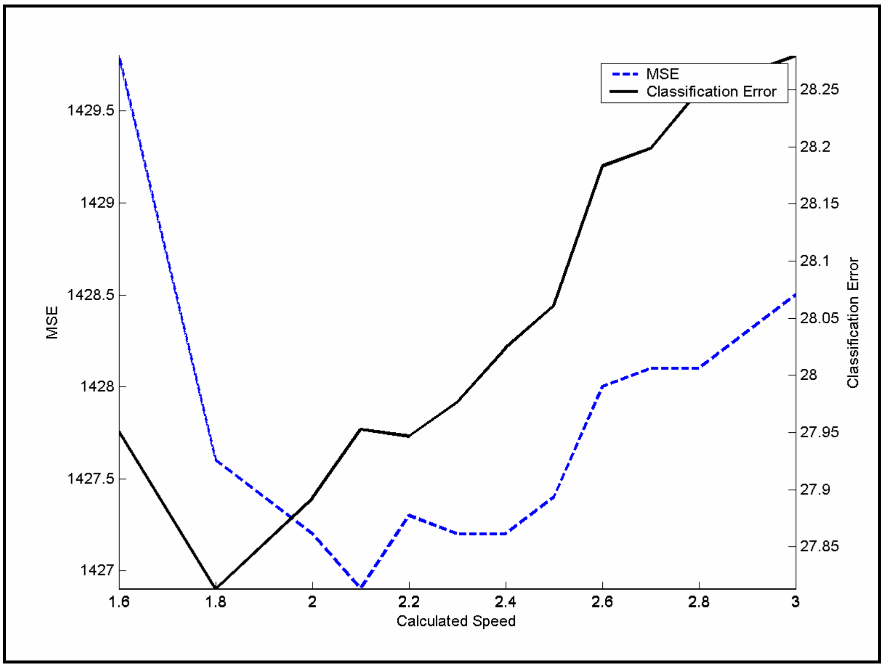


Figure 7.18 Error measures for various upper cut-off trawl speeds (m/s) in the tiger/endeavour prawns trawl signature

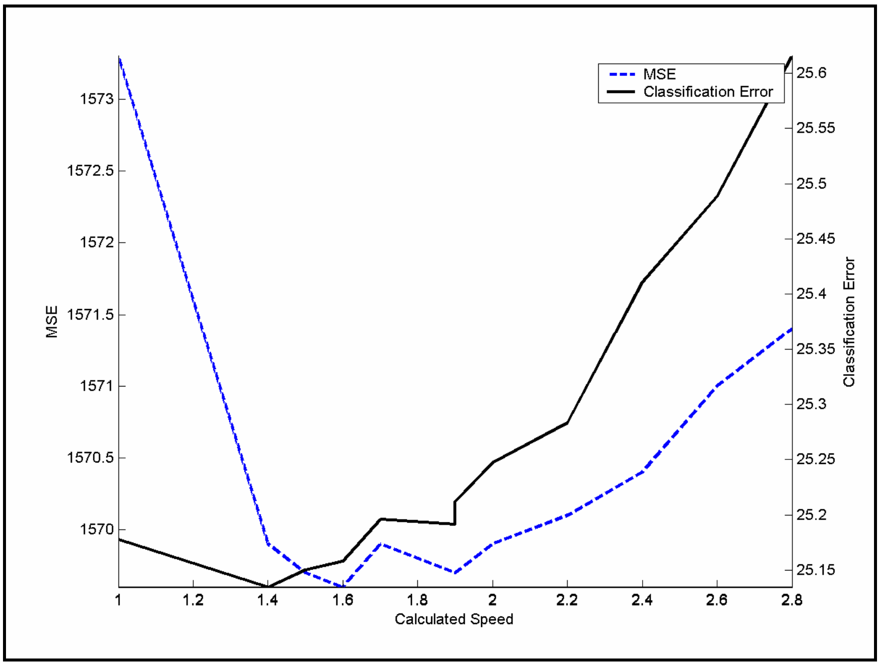


Figure 7.19 Error measures for various upper cut-off trawl speeds (m/s) in the banana prawn trawl signature

7.3.2 *Hourly poll effort maps*

Total effort

Five total effort maps are presented to cover the Queensland coast (Figure 7.20 - Figure 7.25).

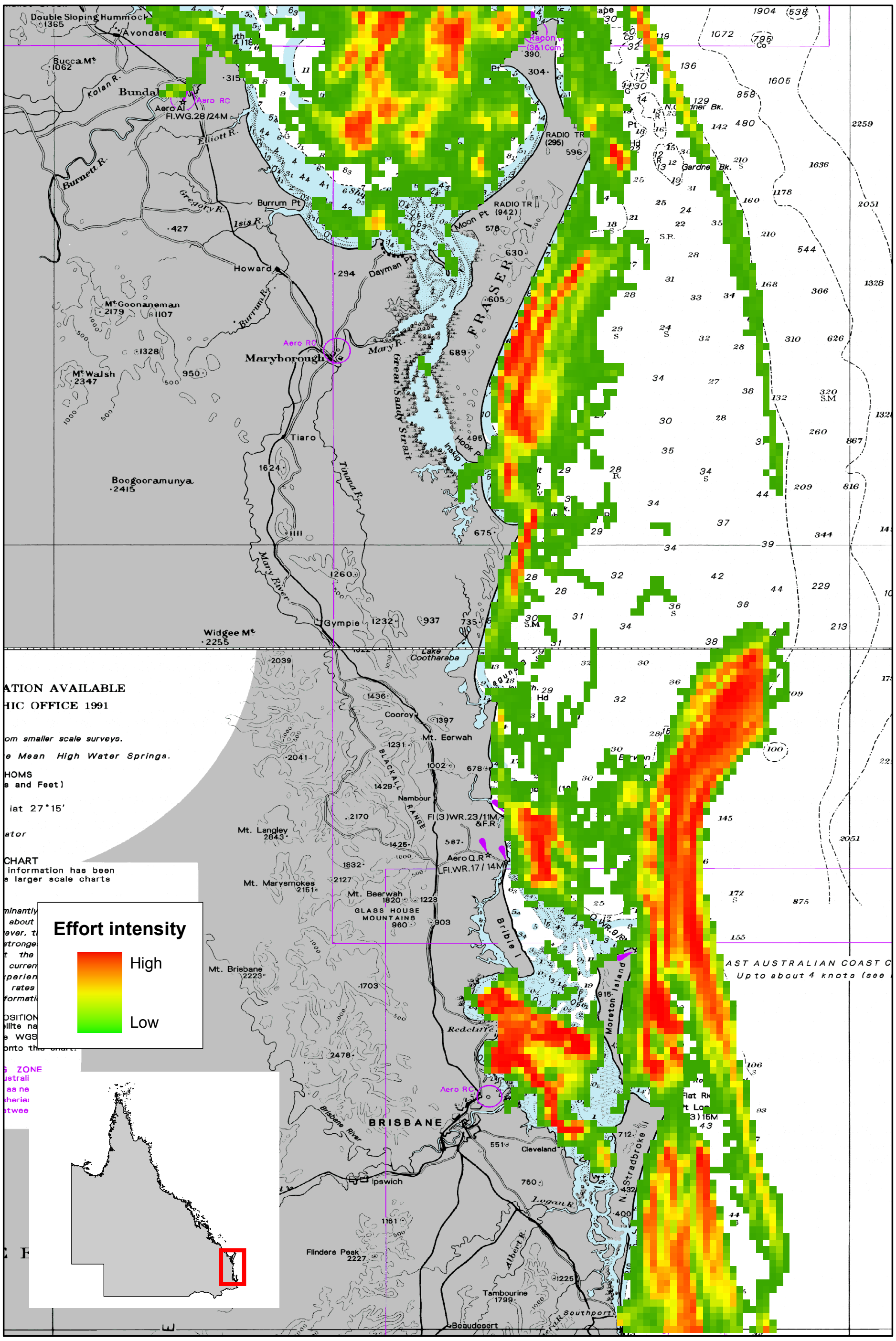


Figure 7.20 Total effort intensity map in far southern section (approximately from 24º40'00''S, 151º55'00''E to 27º55'00''S, 154º50'00''E) at one-minute scale.

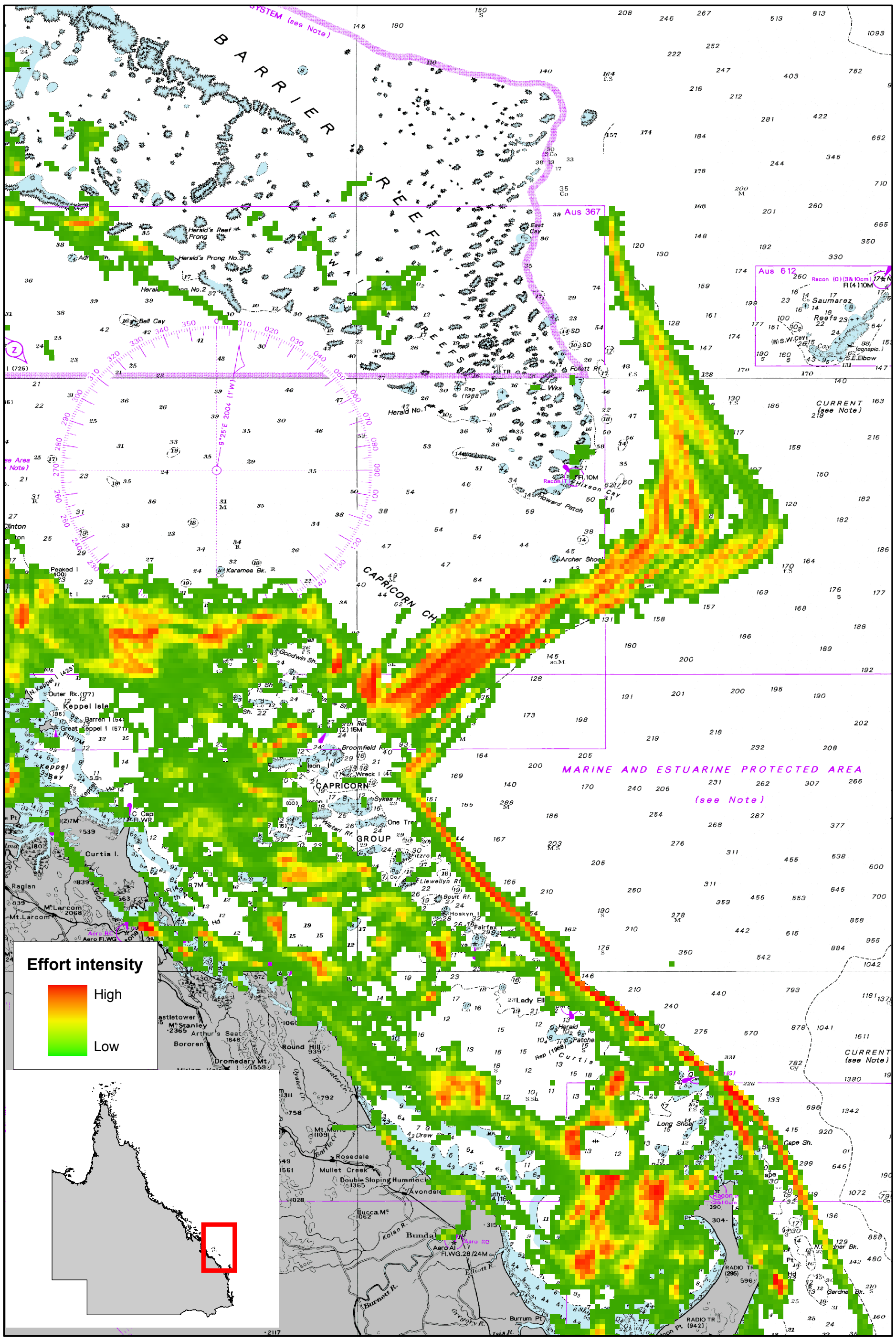


Figure 7.21 Total effort intensity map in mid southern section (approximately from 20º40’00‘’S, 150º50’00‘’E to 25º15’00‘’S, 153º55’00‘’E) at one-minute scale.

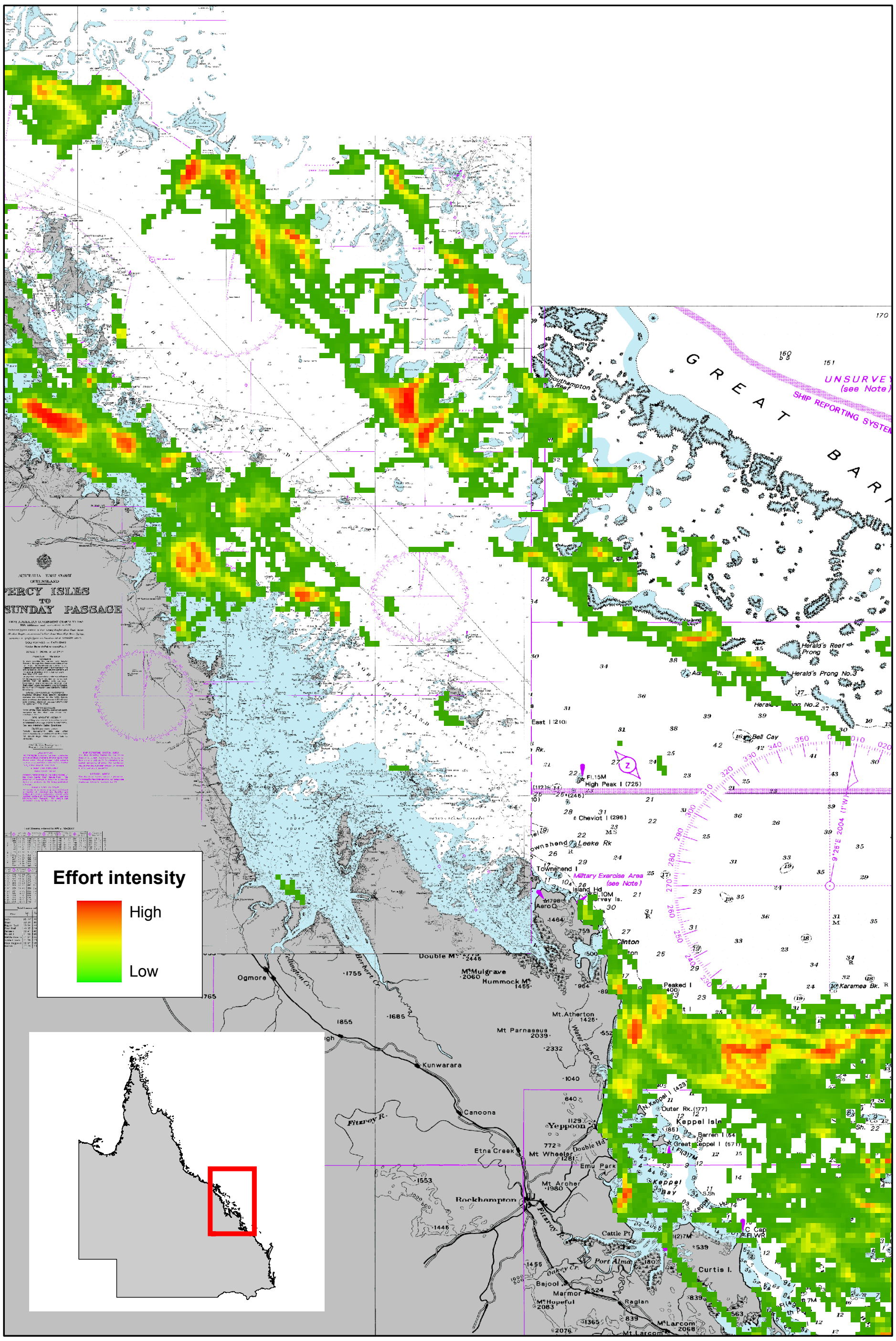


Figure 7.22 Total effort intensity map (approximately from 19º20’00’’S, 148º45’00’’E to 23º55’00’’S, 151º50’00’’E) in mid section at one-minute scale

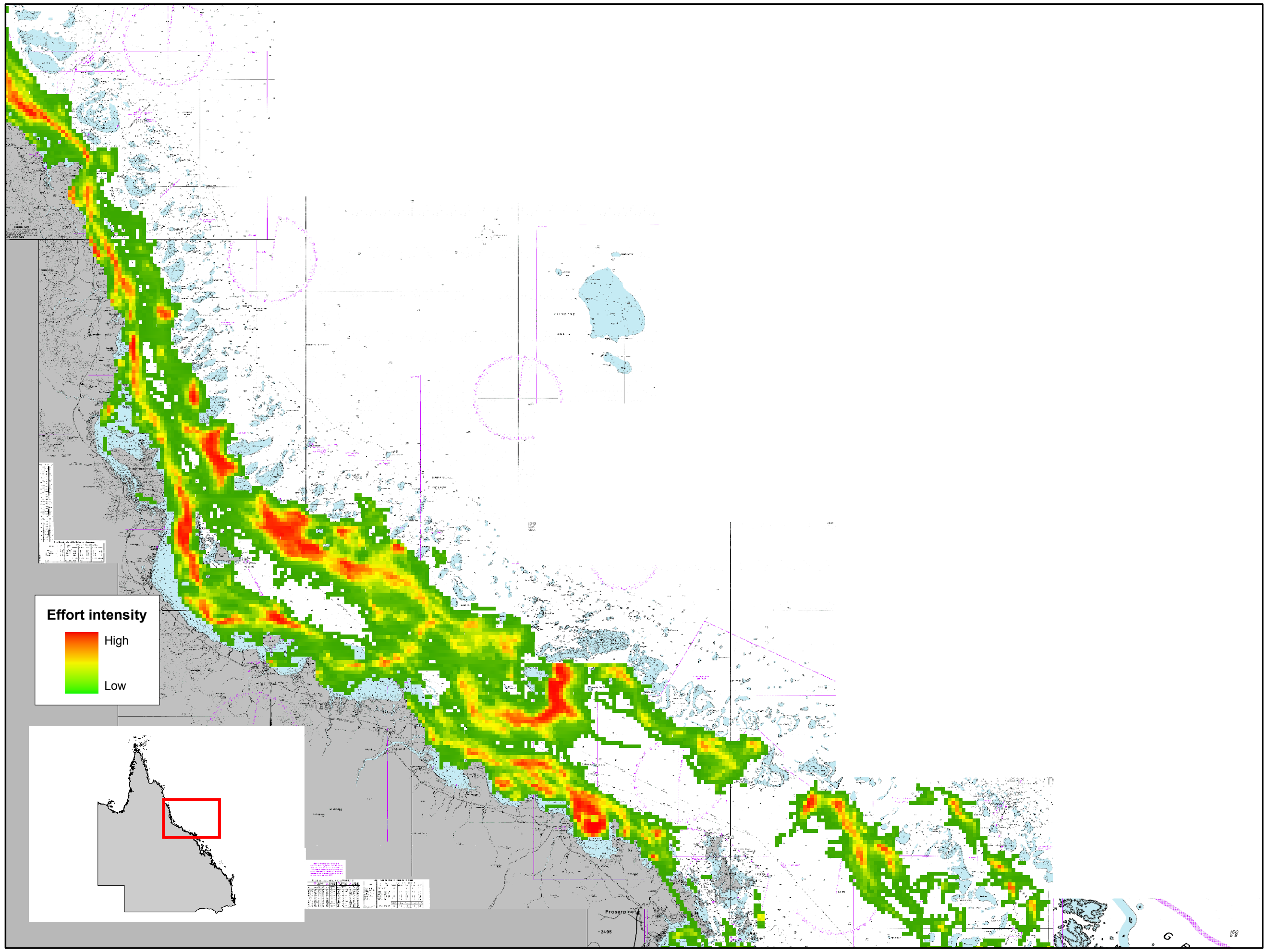


Figure 7.23 Total effort intensity map in the northern section of Queensland (approximately from 16º10’00’’S, 145º30’00’’E to 20º30’00’’S, 151º30’00’’E) at one-minute scale.

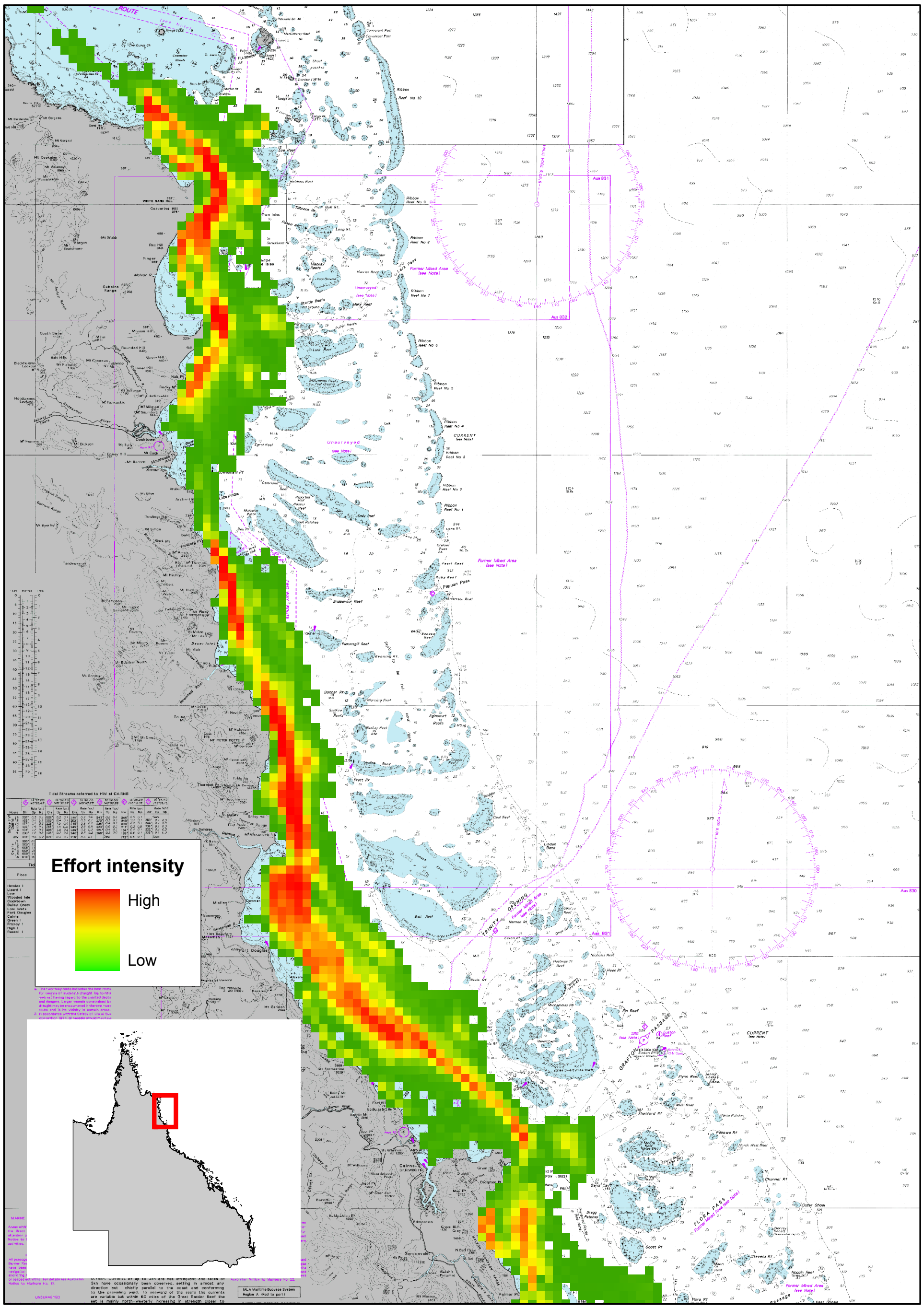


Figure 7.24 Total effort intensity map in the far northern section of Queensland (approximately from 14º30’00’’S, 144º55’00’’E to 17º10’00’’S, 146º45’00’’E) at one-minute scale

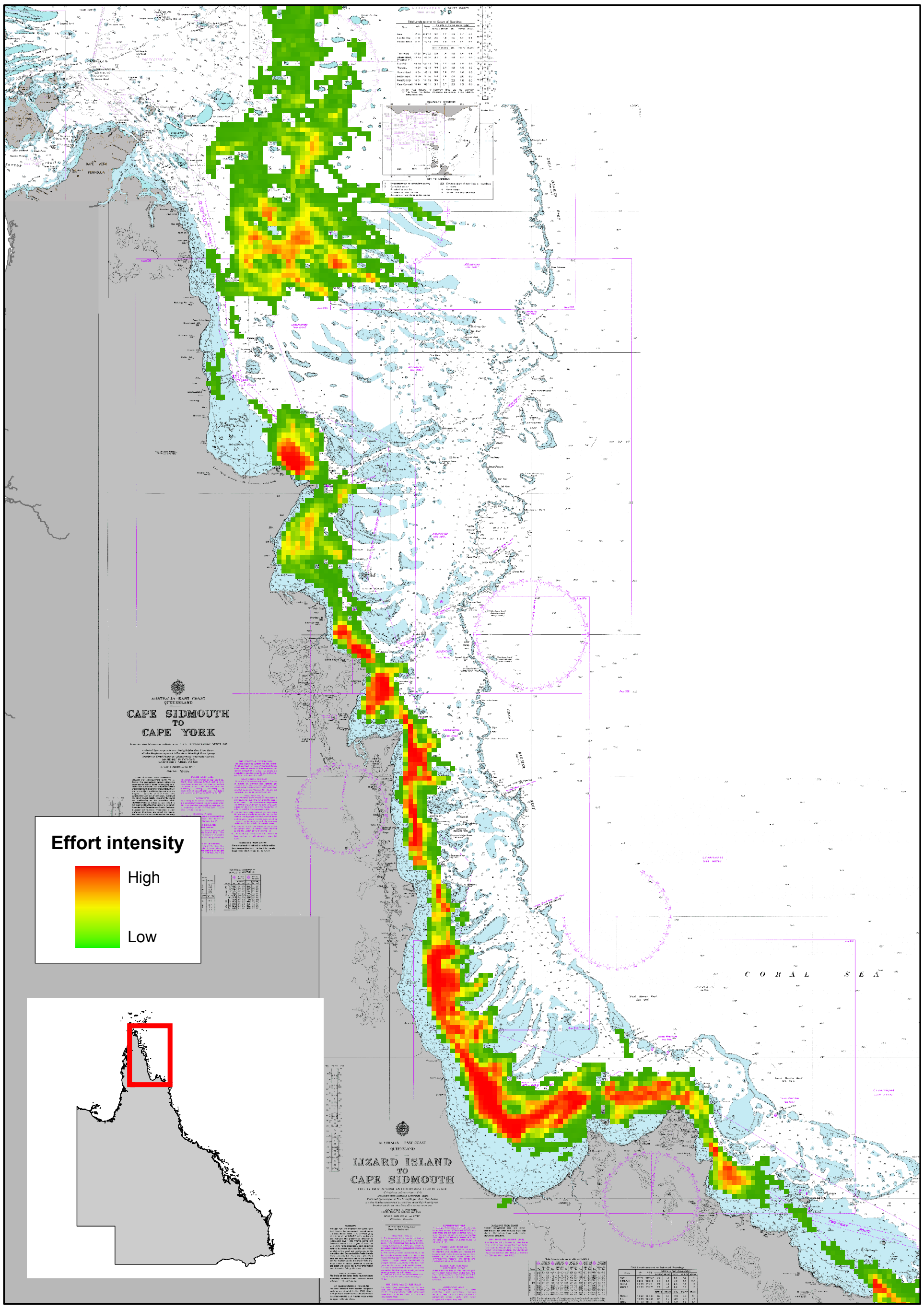


Figure 7.25 Total effort intensity map in the extreme far northern section of Queensland (approximately from 10º10’00’’S, 142º10’00’’E to 14º50’00’’S, 145º20’00’’E) at one-minute scale

EKP

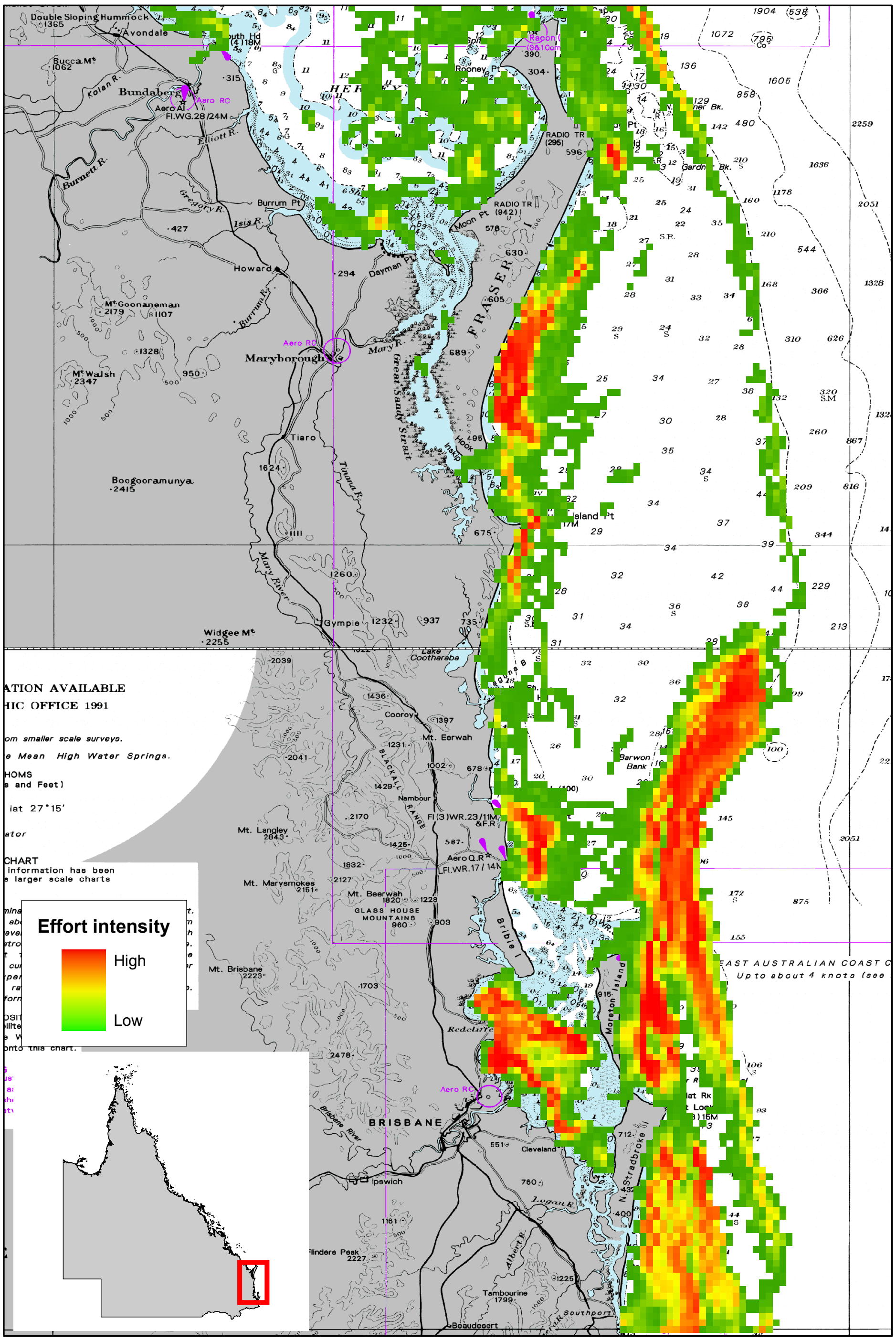


Figure 7.26 Shallow water EKP effort intensity at one-minute scale maps (approximately from 24º40’00’’S, 151º55’00’’E to 27º55’00’’S, 154º50’00’’E).

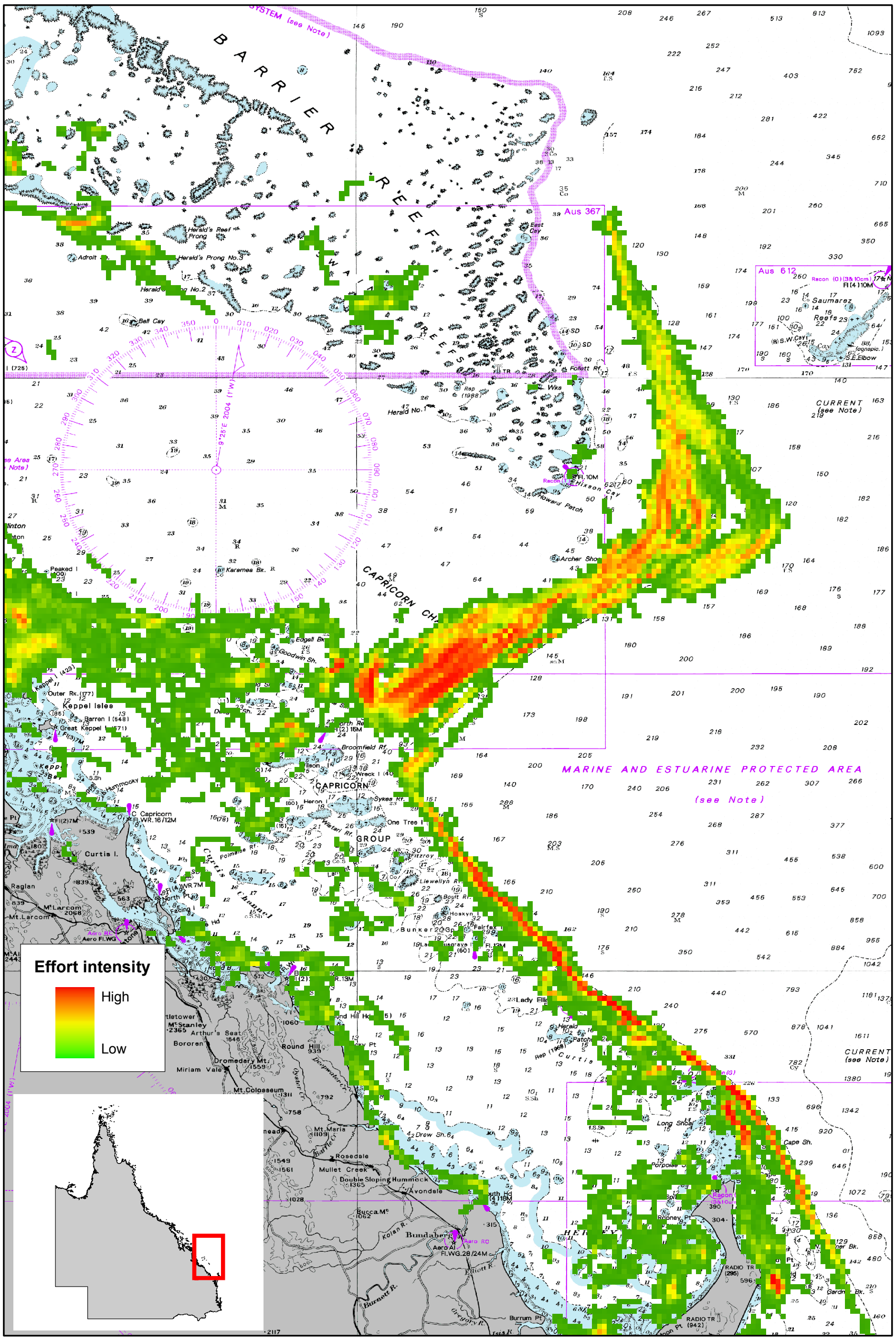


Figure 7.27 Deep water EKP effort intensity map at one-minute scale (approximately from 20º40’00’’S, 150º50’00’’E to 25º15’00’’S, 153º55’00’’E).

Scallop

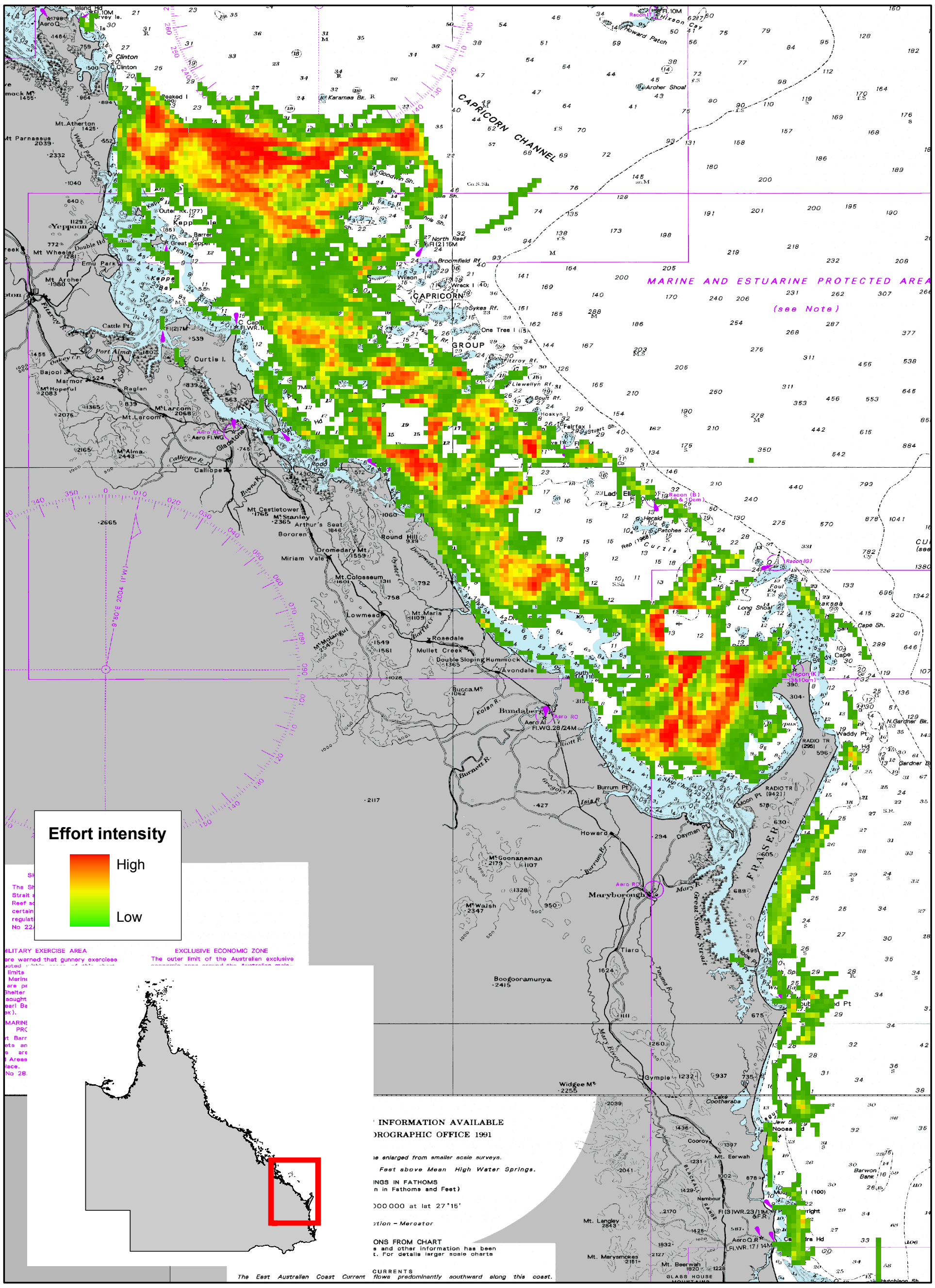


Figure 7.28 Scallop effort intensity map at one-minute scale (approximately from 22º20’00’’S, 150º30’00’’E to 26º55’00’’S, 153º40’00’’E)

Tiger/Endeavour Prawns

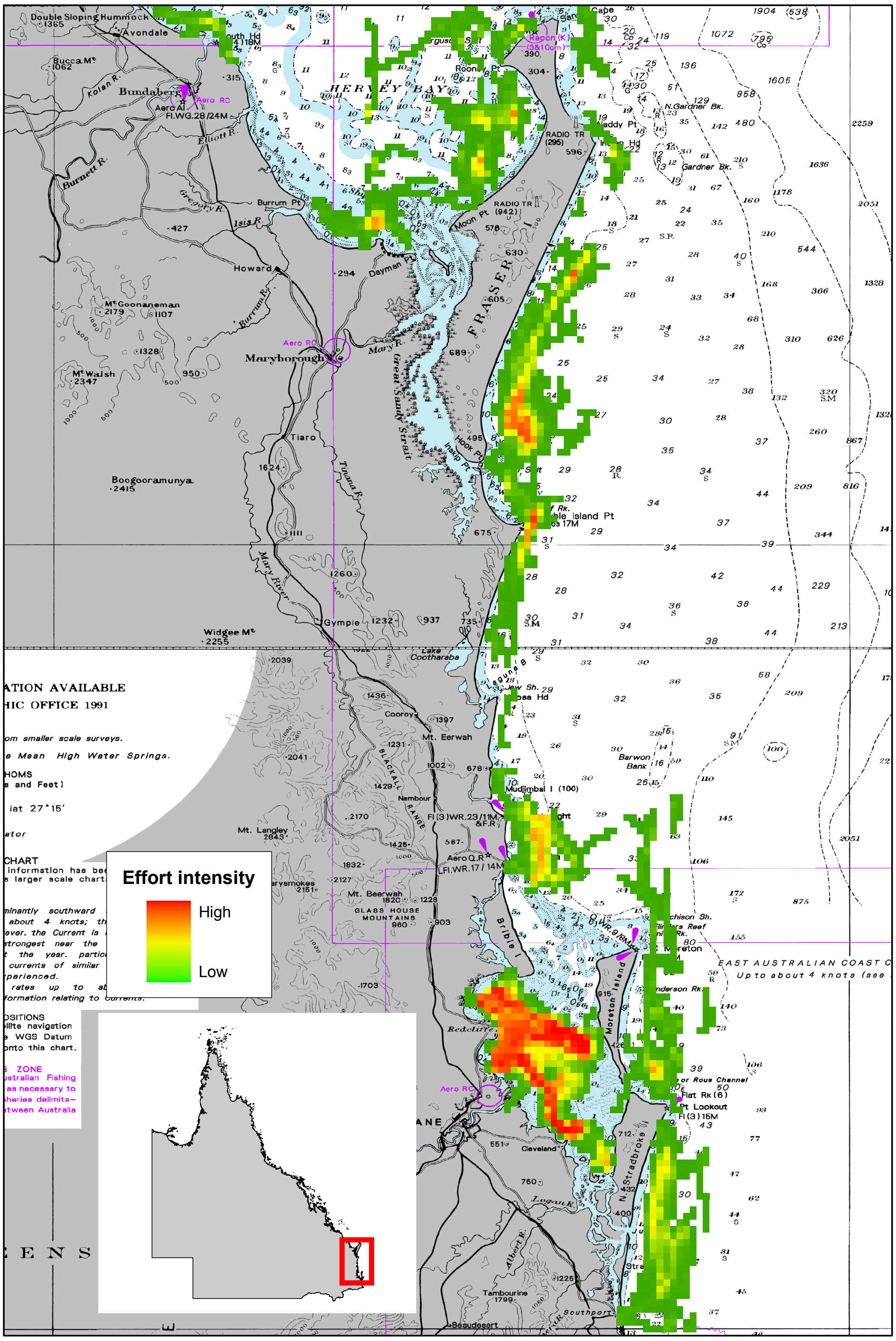


Figure 7.29 Tiger/endeavour prawns effort intensity map in far southern section of Queensland (approximately from 24º40’00’’S, 151º55’00’’E to 27º55’00’’S, 154º50’00’’E) at one-minute scale.

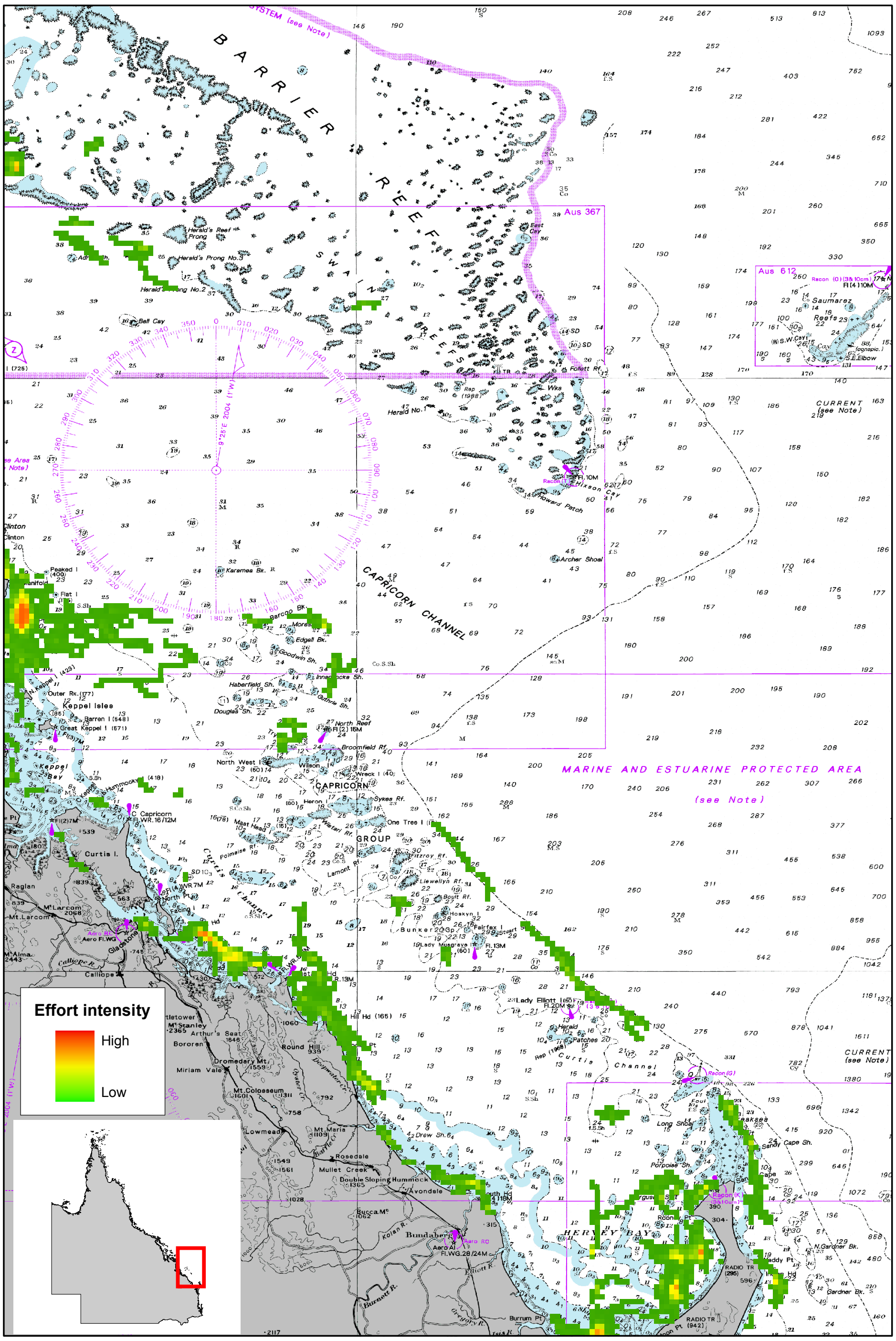


Figure 7.30 Tiger/endeavour prawns effort intensity map in mid southern section of Queensland (approximately from 20º40’00’’S, 150º50’00’’E to 25º15’00’’S, 153º55’00’’E) at one-minute scale.

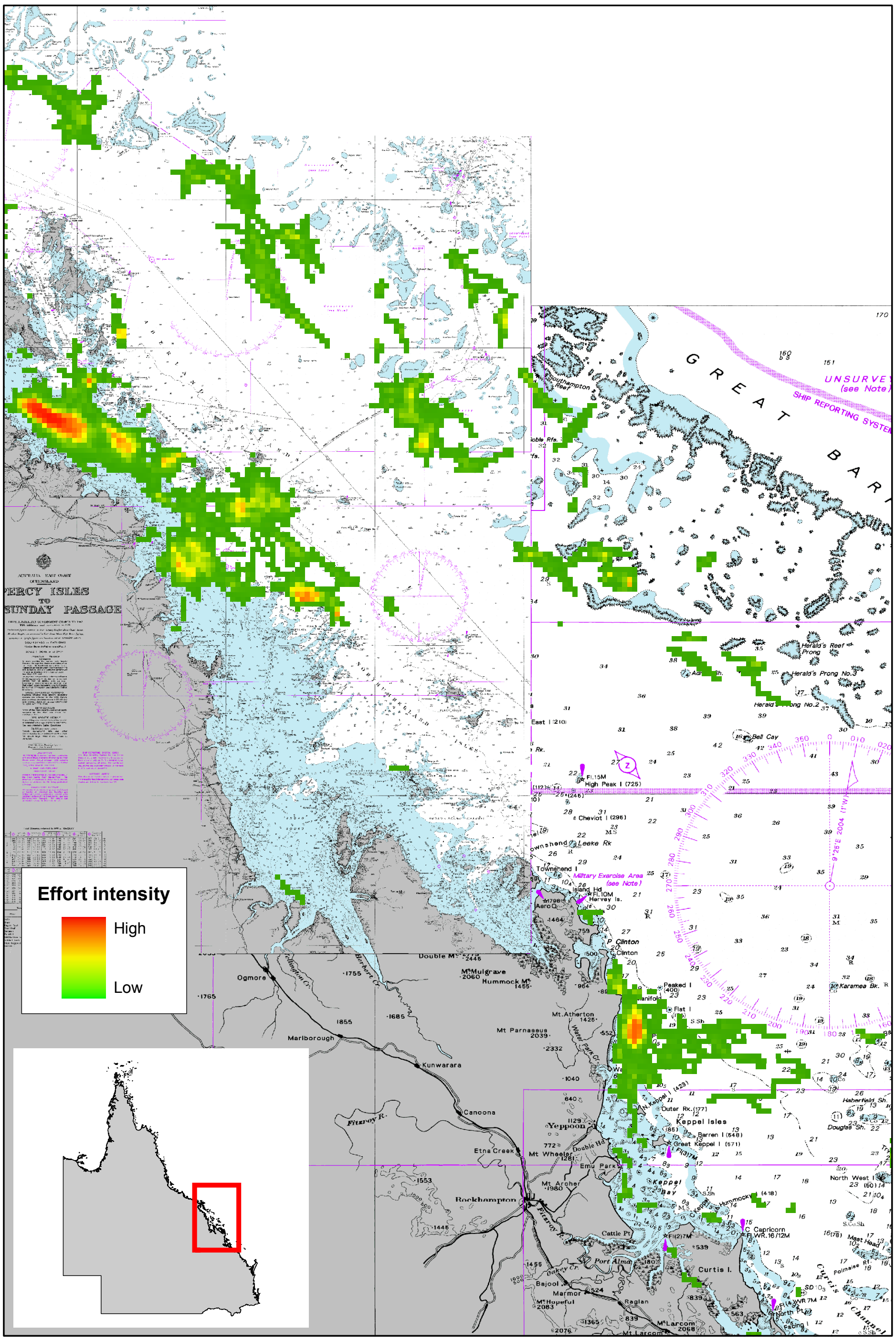


Figure 7.31 Tiger/endeavour prawns effort intensity map in mid section of Queensland (approximately from 19º20’00’’S, 148º45’00’’E to 23º55’00’’S, 151º50’00’’E) at one minute scale.

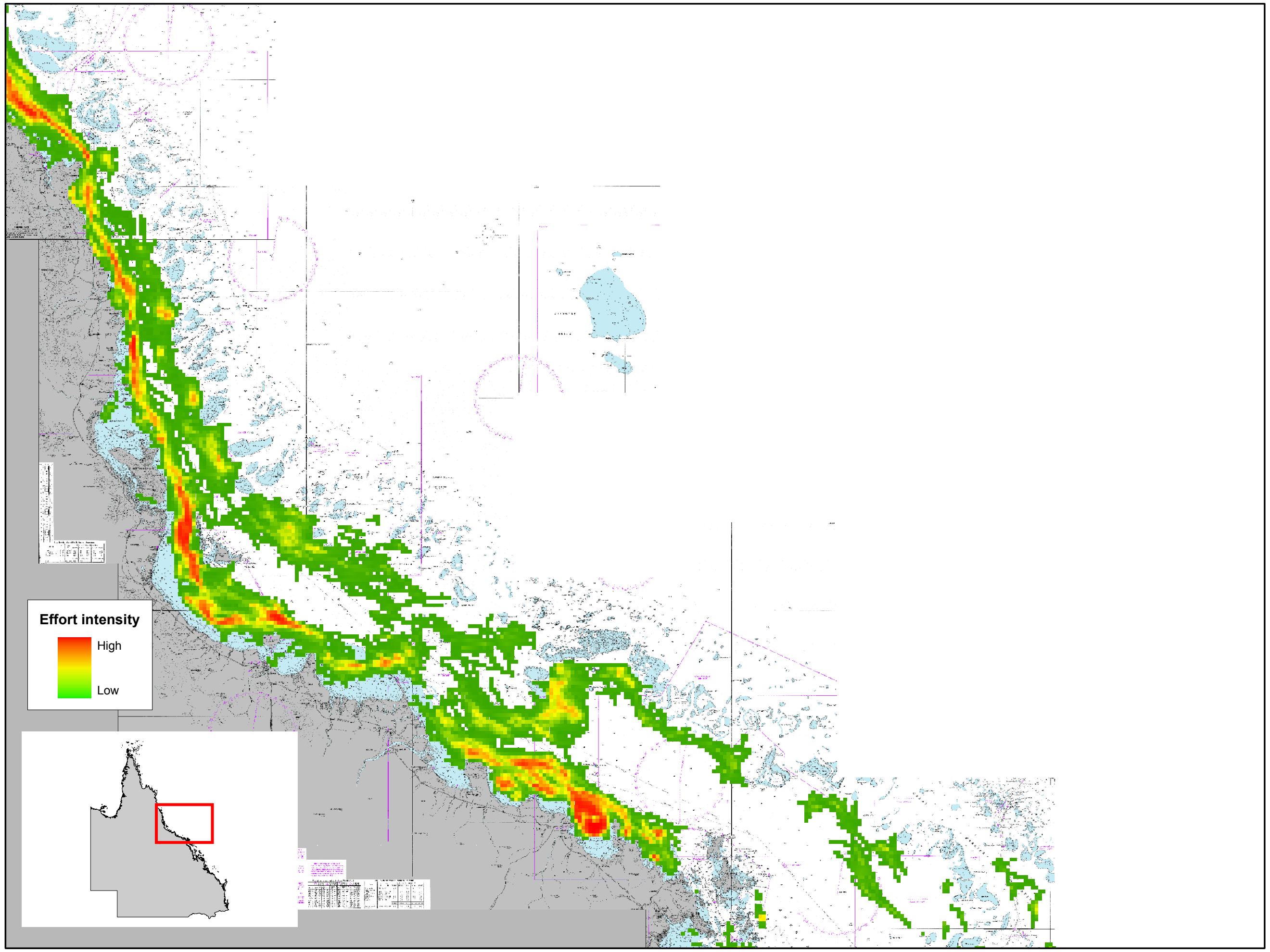


Figure 7.32 Tiger/endeavour prawns effort intensity map in the northern section of Queensland (approximately from 16º10'00''S, 145º30'00''E to 20º30'00''S, 151º30'00''E) at one-minute scale.

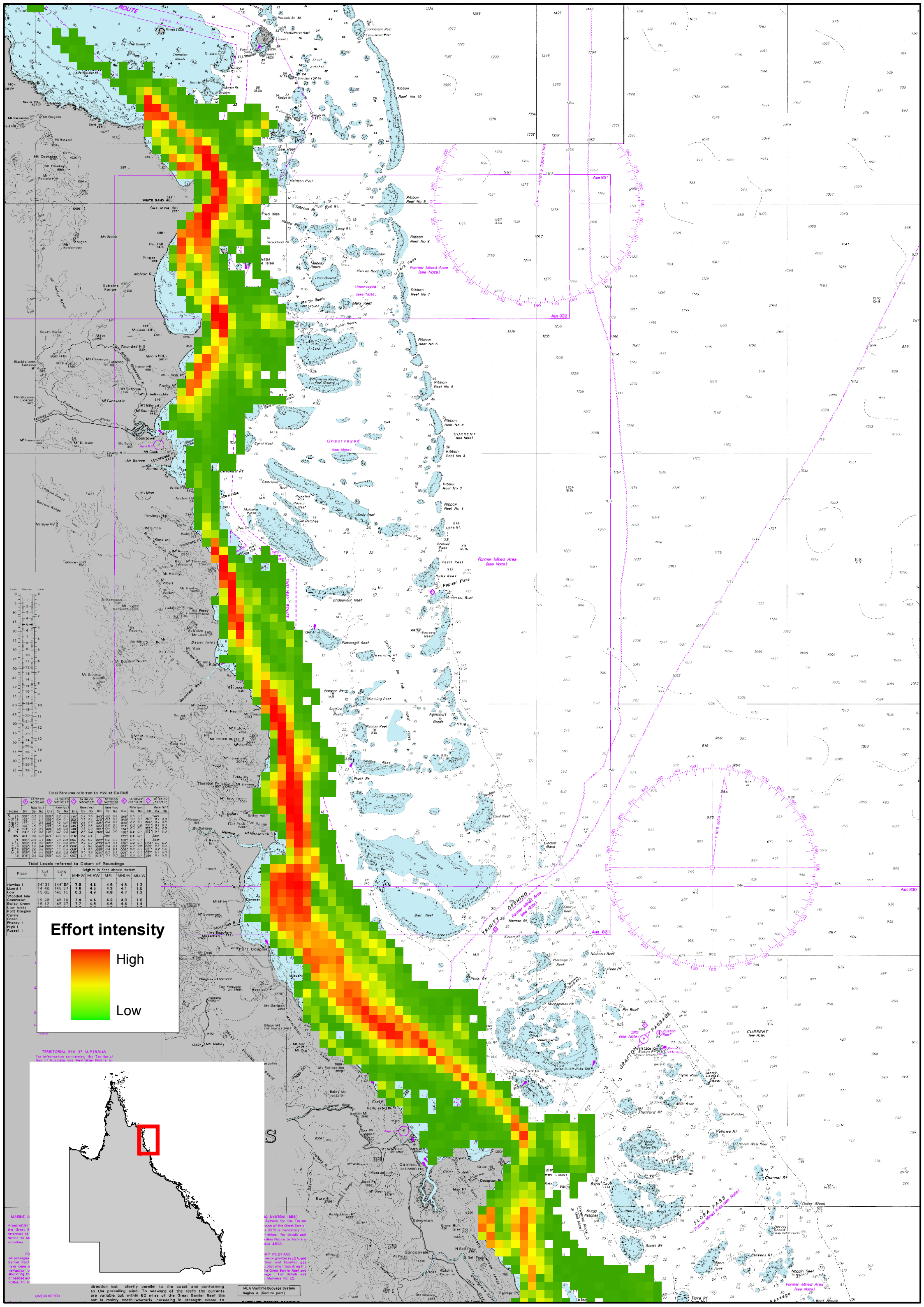


Figure 7.33 Tiger/endeavour effort intensity map in the far northern section of Queensland (approximately from 14º30'00''S, 144º55'00''E to 17º10'00''S, 146º45'00''E) at one-minute scale.

7.3.3 *Comparison of spatial resolution*

To demonstrate the resolution of VMS data, Figure 7.34 shows the 6 x 6 nm^2 grids of effort with the underlying VMS trawling data superimposed in black. Note that the data have been transformed to provide confidentiality. It clearly shows the level of aggregation evident in this particular fishery. It also illustrates the benefit of mapping at finer spatial scales.

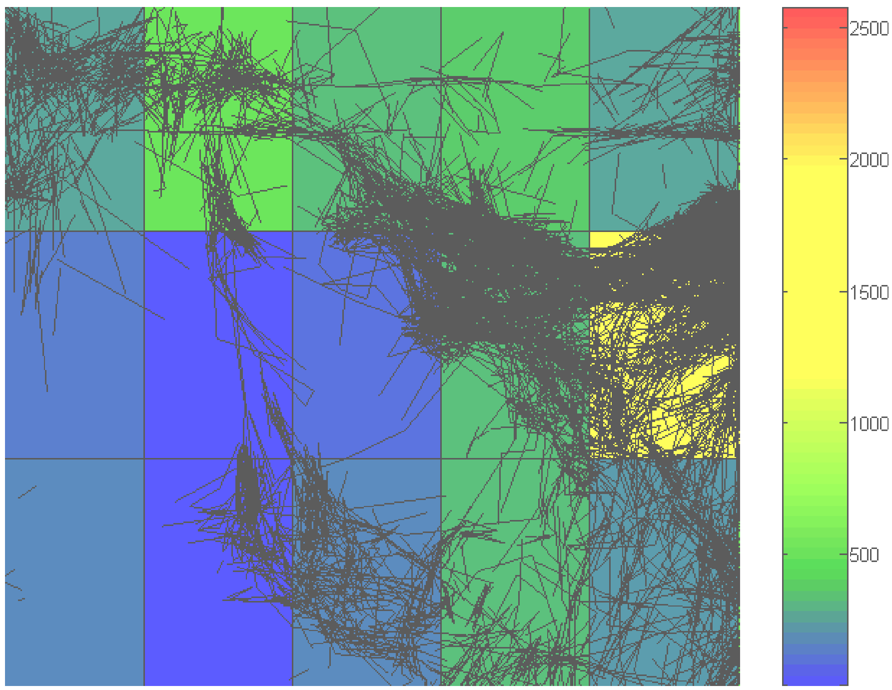


Figure 7.34 Comparison of resolution of VMS track information and effort based on 6 x 6 nm^2 grids.

7.4 Discussion

The development of trawl signatures for selected fisheries in Queensland is the first stage in a process for improving the accuracy of stock assessments. The ability to map effort consistently at a finer scale should substantially improve estimates of fishing intensity, underlying resource density and other stock assessment parameters previously difficult to quantify. Improved parameter estimation and spatial resolution will give management more information when deciding the regulation of effort consistent with sustainability. In the case of implementing closures potential conflict with commercial fishers by closing non-critical areas is significantly reduced.

We initially proposed to fit a normal mixture model to these data to categorise trawling and steaming. However, closer examination of these histograms causes some concern as the modes corresponding to trawling and steaming do not seem to follow a normal distribution. In particular in the tiger/endeavour and banana prawn fisheries, the histograms have an extra mode or peak as the calculated speed approaches zero. At first this was perplexing because the other fisheries, although showing some extra density in this region, did not exhibit these peaks. However, in histograms based on the small amount of high-frequency polling data, the extra peaks were not evident. Therefore it would seem that the extra peak is an artefact of the discrete polling.

Another issue with the speed-based rule in a mixture context is that a cut-off speed corresponds to the point at which probability of trawling is 50%. This may not be appropriate and we may want to either use more stringent criteria (e.g. we must be 80% sure the vessel is trawling). An alternative approach would be use a probabilistic allocation of trawling activities rather than a fixed cut-off. This probabilistic approach will be explored in the next chapter.

To examine the effect of polling let us assume that the optimal actual trawl speed for a given fishery does not vary considerably. Now consider we poll every five seconds then the resulting histogram would be as in Figure 7.35.

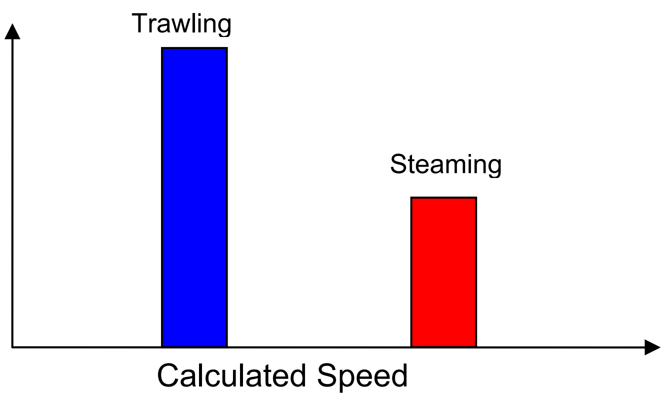


Figure 7.35 Example showing high-frequency polling of optimal trawl/steaming speeds

Now let us make our polling frequency longer, for example, every hour. This will produce two components of variation or error. Firstly, vessels might not spend the whole hour either trawling or steaming – they might do both. This will alter the calculated speed which is taken over the hour, by biasing the trawling speeds upward and the steaming speed downward, for example, see Figure 7.36. In a similar way vessels might slow down and stop during the polling interval.

The second source of bias is the fact that vessels do not always travel in straight lines. Since the calculated speeds are obtained by basically connecting the two VMS polls by a straight line, the calculated speed might be biased downward as the vessels have actually travelled a greater distance in the hour than the VMS indicates (see Figure 7.37). Similarly, if the vessel trawls in straight lines but turns back upon itself, and if the lengths of these tows are very short, then the VMS polls will be close together even though the vessel has actually traversed a greater distance.

Considering the presence of the extra mode in the banana and tiger/endeavour prawns and the absence in the EKP and scallop, this would seem to infer that the banana and tiger/endeavour prawns have more curved or shorter trawl tows than the scallop and EKP.

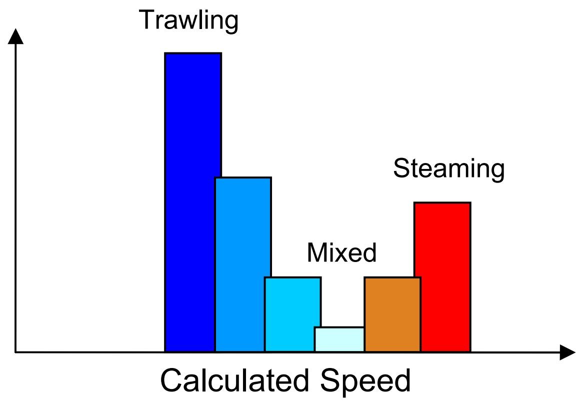


Figure 7.36 Example of the effect of longer polling times introducing mixed trawling/steaming.

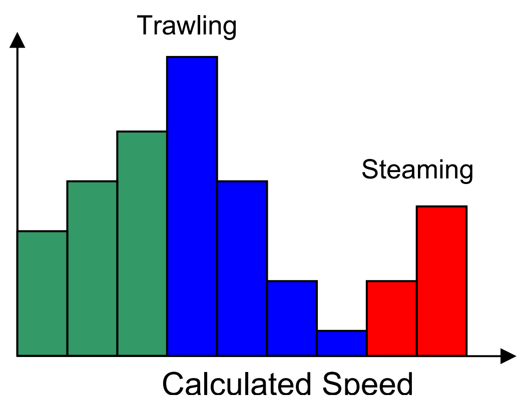


Figure 7.37 Example of the effect of longer polling time in the presence of highly curved or short trawl tows.

Considering the limitations of the mixture model, we considered alternatives to both the model and variables used within the model. Chapter 8 deals with Hidden Markov chain and simple filter methods as alternatives to the use of the mixture model. We have considered using GPS data, high-frequency VMS data, and high resolution logbook effort data as alternative variables in estimating effort. Each model had positives and negatives, the filter method (simple speed cut-off rule) is easiest to implement but it is least discriminating. More-complex models are potentially robust and more accurate, but computationally complex.

The effort maps, however, clearly showed the spatial coverage. Mapping effort at such a fine spatial scale will allow more equitable decisions to be made in the spatial management of the fisheries.

8 REFINED TRAWL SIGNATURE AND TRAWL TRACK DEFINITIONS USING A HIDDEN MARKOV MODEL APPROACH

David Peel and Norm Good

8.1 Introduction

VMS is being adopted as compliance tools in many fisheries (Smith 2001). A secondary outcome of this technology is a large database of fine-scale spatial and temporal information on vessel positions. These data can be invaluable in stock assessment, research and management in a number of ways. Some examples of the uses of fine-scale spatial data include:

- simple mapping of fishing intensity at a fine temporal and spatial scale (Afonso-Dias 2000; Gribble, Peel *et al.* 2006; Larcombe, McLoughlin *et al.* 2001; Marrs, Tuck *et al.* 2002; Mejias. 1999; Murawski, Wingley *et al.* 2005)
- measuring depletion (Deng, Dichmont *et al.* 2005; Gedamke, DuPaul *et al.* 2004)
- aiding in assessing trawling impacts (Rijnsdorp, Buys *et al.* 1998; Trimmer, Petersen *et al.* 2005) or
- estimating spatial resource abundance (Bertrand, Burgos *et al.* 2005; Good and Peel 2005; Vignaux, Vignaux *et al.* 1998).

There are not many full-scale studies of the use of VMS for scientific research in the literature. This may be because much of the work of this type is done in-house and not suitable for publication (Nishida and Booth 2001).

VMS collects positional data of vessels at regular time intervals (i.e. a poll). This polling occurs irrespective of the vessel's activity (i.e. trawling, at anchor, or steaming). However, to use the data in a meaningful way we require an indication of vessel activity. In general we are mainly interested in trawling activity, although it is foreseeable that in some instances the spatial distribution of steaming or the location of anchorages may be of interest. In some VMS implementations vessel activity may be known, such as when information from winch sensors is available (for example, Mejias 1999). However, since this is generally not the case, some method to estimate vessel activity is required.

As there is a physical limitation on how fast a vessel can draw a trawl net through the water (which is considerably lower than normal vessel steaming speed) an obvious indicator of vessel activity is the vessel speed. There would, in theory, be a 'maximum optimal trawl speed' and any vessels moving faster than this could be deemed to be steaming. However, in reality this is not the case as the trawling speed at any given time is determined by a number of factors, for example: vessel engine power, currents/tides, weather, size of net/type of gear, target species, captain/crew preference, frequency of tows. Also, VMS data do not directly provide a measure of vessel speed over the poll period but rather a 'calculated' vessel speed. The calculated vessel speed is simply based on the direct distance from the previous poll divided by the polling time interval. Obviously, vessels do not always travel in straight paths, so this measure of speed will generally be biased downwards. Since we would expect the vessel to take straighter more direct routes when steaming than when trawling (depending on the fishery), this bias might actually help discriminate between trawling and steaming. The calculated speed is also affected by non-trawling activity within the polling period (e.g. bringing in/letting out/emptying nets, turning, sorting catch) during which time the vessel may slow.

The second indicator of vessel activity is the time of day. In many fisheries the fishers trawl only during certain periods of the day (e.g. day or night) due to regulation or target species behaviour. In the simplest case, all points occurring during the non-trawling period could be removed. However, after examining the data and consulting with fishers it seems that there is some variation (particularly at the boundary between night and day) on when fishing occurs.

The existing literature does not seem to contain any detailed work about determining vessel activity from VMS data. It must be assumed that in practice some rudimentary rule on speed is used to filter out vessels travelling too fast to be trawling. However, this approach has several flaws, or limitations:

- It ignores the temporal correlation of the data; that is, the probability that a vessel is trawling at the next poll is not independent of the activity at the current poll.
- The cut-off speed at which vessels are filtered is sometimes chosen with no statistical rational.
- An outright classification of the data, into vessel activity, does not provide any measure of the uncertainty of the classification, e.g. polls that correspond to speeds at the boundary of plausible fishing speeds are less likely to be completely fishing.

If a calculated speed cut-off is to be used to determine vessel activity, it would be advisable to use the data to determine the cut-off, preferably allowing different cut-offs for each vessel and fishery.

Initially we investigated an approach using finite mixture models (McLachlan and Peel 2000) fitted to individual vessels historical speed and time of day data to determine vessel activity. This approach did solve many of the issues, however, the temporal auto-correlation of the data was not addressed. To model the temporal correlation a natural avenue to pursue was Hidden Markov Models (which can be thought of in a mixture model framework as a mixture model with auto-correlated hidden membership variables).

8.2 Method

8.2.1 *Description of model*

To provide a probabilistic determination of vessel activity a discrete Hidden Markov Model (HMM) was used. HMMs are being used in many applications including speech recognition, econometrics, biology, and image processing (see MacDonald and Zucchini 1997 ,or McLachlan and Peel 2000 for a more complete list of references). A HMM is a statistical model of a system that follows a markovian process. That is, the conditional probability of the future states given the present and past state depends only on the present state.

In the context of this chapter the model for a particular vessel consists of a number of states, or nodes, corresponding to the vessel's activity. Each of these nodes is connected by predefined pathways. The process involves the vessel discretely moving around this network with each step corresponding to a polling event. Some pathways loop back to the same state allowing a possible step to be that the vessel does not change state or activity.

In practice, we do not observe the true vessel activity (i.e. the current state) but rather a quantity – the distribution of which is dependent on vessel activity and therefore is indicative of the current state. In this application an obvious indicator of state or activity is the vessels calculated speed. So the distribution of calculated speed is assumed to be dependent on which state the vessel is in at that time. A graphical representation of the model we used is given in Figure 8.1. The model includes the obvious nodes corresponding to trawling, steaming and at anchor (stationary) as well as special 'Entry' and 'Exit' states.

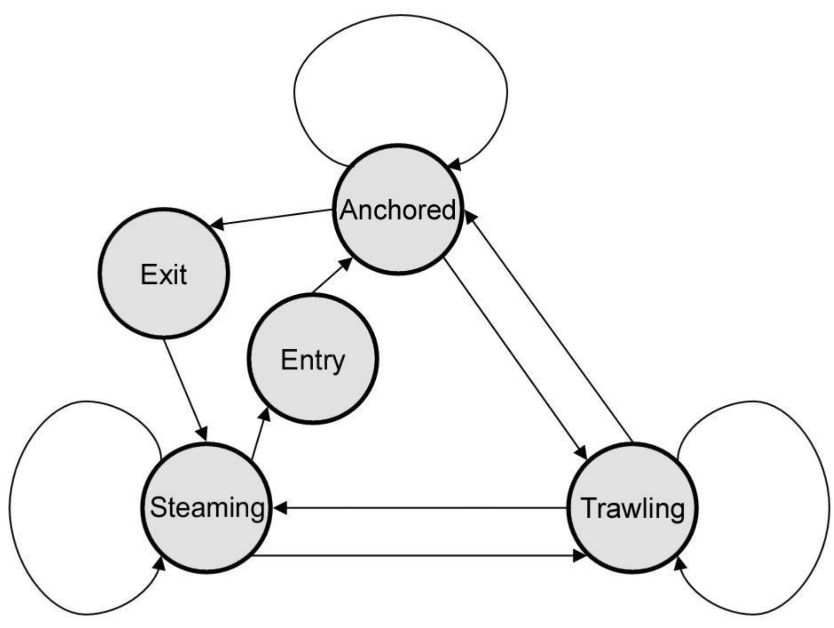


Figure 8.1 Representation of a Hidden Markov chain with transition pathways indicated.

The purpose of the Entry and Exit nodes is to address an important issue in the data. Generally polls corresponding to vessels at anchor can be easily identified, as the calculated speed is very close to zero. However, due to discrete polling, when vessels approach or leave an anchorage at steaming speed, a vessel will often be misclassified as trawling. To illustrate why this occurs, suppose a vessel is polled while returning to port at steaming speed, then arrives in port and is stationary, until it is polled again. The calculated speed for this interval will be greater than zero as the vessel was steaming for the first portion of the poll interval, but will be less the steaming speed (i.e. closer to trawling speed) due to the vessel being stationary during the second portion of the polling interval. The Entry/Exit nodes model this transition between anchorage and steaming. Similar nodes were not added between anchoring and trawling as from discussion with fishers, it is evident that trawling may occur when entering and exiting the anchorage. This was mainly the case when anchoring was very close to the actual trawl ground.

Each vessel in the fleet is modelled independently in this way. This allows for individual differences in behaviour between vessels due to any number of unknown factors, e.g. vessel size, engine power or location.

To describe the HMM more formally, consider a Markov chain with S states. The state of the process at time t is denoted by an indicator variable z_{tj}, where $z_{tj} = 1$ if the process is in state j at time t and equals 0 otherwise. The transition probability matrix P is defined such that:

$$\Pr[z_{t+1,i} = 1 \mid z_{t,j} = 1] = P_{ij}$$

The current state at any given time t is not observed directly so z is a hidden variable. However, the measurement of calculated speed y_t is observed and it is assumed to be distributed as:

$$f(y_t \mid z_t; \phi) = \prod_{j=1}^{S} f_j(y_t, \theta_j)^{z_{tj}}$$

where $\phi = \{\theta_1, \theta_2, ... \theta_S\}$ and θ_j denotes the unknown parameters in the density (j=1,… S).

In our model, as seen in Figure 8.1, the number of states S is taken to be 5 and the state densities f_j (j = 1,… S) were taken to be

$$f_1\left(y;\theta_1\right) \sim HalfN(y;0,0.01) \qquad \text{(Anchorage)}$$
$$f_2\left(y;\theta_2\right) \sim \sum_{k=1}^{g} \pi_k N\left(y;\mu_{2,k},\sigma_{2,k}^2\right) \qquad \text{(Trawling)}$$
$$f_3\left(y;\theta_3\right) \sim U\left(y;0,\mu_5\right) \qquad \text{Entry}$$
$$f_4\left(y;\theta_4\right) \sim U\left(y;0,\mu_5\right) \qquad \text{Exit}$$
$$f_5\left(y;\theta_5\right) \sim N\left(y;\mu_5,\sigma_5^2\right) \qquad \text{(Steaming)}$$

where $N(y;\mu,\sigma^2)$ and $HalfN(y;\mu,\sigma^2)$ denote the normal and half-normal distribution, respectively, with means μ and variances σ^2, and $U\left(y;a,b\right)$ denotes the Uniform distribution with bounds a and b. The parameters π_k, μ_i and σ_i (i = 1,… 5; k = 1,… g) are the unknown parameters to be estimated.

The choice of state distributions was based on examination of the empirical distributions of VMS-calculated speeds stratified by an approximate outright classification. It was found that the distribution of trawl speeds was often almost bimodal for some fisheries. Upon further investigation this seemed less prevalent when the polling frequency was increased. So the bimodal distribution is possibly due to non-trawling activities occurring during the polling period (e.g. sorting catch or turning around). To handle the non-normality of the data a 'g component' finite mixture model was used with unequal variances (for the example application in this chapter g = 2 was used). The entry/exit distributions were assumed to be a uniform distribution with limits taken as the means of the anchorage and steaming distributions. The anchorage distribution was taken as a half-normal fixed with mean zero with a small variance to encompass speeds very close to zero.

To model the effect of the time of day, the transition matrix P was taken to depend on time of day d. In other words, we are assuming that the probability of entering or moving to a state will depend on the time of day. For example, in a particular fishery the probability of moving to the trawling state during the day may be very low and anchorage/steaming transition probabilities may be high. During the night the opposite will occur. It is assumed that the vessel's trawl and steaming speeds will be unrelated to time of day, i.e. the state distribution parameters ϕ do not depend on time of day.

The model was also extended to model the change in behaviour of vessels corresponding to the fishery they were targeting. This was done by allowing the transition probabilities to be dependent on fishery c. For example, in some fisheries where very long trawls are common, the transition probability of remaining in the trawling state would be higher than a fishery where short quick trawling is the norm. In addition, since it appears vessels may trawl at different speeds when in different fisheries, the state distribution was taken to depend on fishery $\phi_c = \left\{\theta_{1c},\theta_{2c},...\theta_{Sc}\right\}$.

So for time of day d and fishery c the transition matrix P was taken to be of the form:

$$P_{ijdc} = \Pr[z_{t+1,i} = 1 \mid z_{t,j} = 1 \ \& \ D = d \ \& \ C = c]$$

For ease of use, we shall denote P_{ijdc} as $P_{ij}\left(d, c\right)$.

The fishery c can be determined by the spatial location, the time of day and caught species. It should be pointed out that even though we state c is related to fishery, the aim is not necessarily to accurately identify true targeted species. Rather, we simply want to provide a mechanism in the model to handle heterogeneity caused by differing trawl behaviour. So it is not critical to correctly identify target species as long as fishing behaviour within our defined 'fishery' is reasonably homogenous for the vessel in question. However, it does seem that target species is a reasonable criterion on which to base the grouping strategy. Furthermore, the fishery definition framework, if required, could be used to model other extra heterogeneity in the data. For example, a single vessel which had different skippers fished differently over time.

In our application, information on catch is available to determine fishery. If this was not the case, then an alternative approach could be to allow the model to also estimate target species/fishery. Assuming that the change over of fishing gear required to switch between fisheries only occurs while a vessel is anchored then the model can be extended as per Figure 8.2. In this model each fishery has its own corresponding branch of the Markov model, with the interchange node being a common anchorage node.

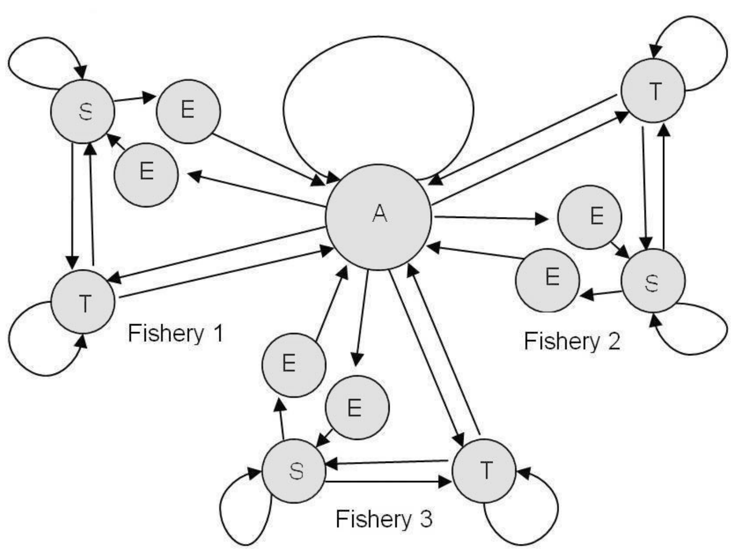


Figure 8.2 Example of HMM structure that could be used to model greater than one fishery in the absence of reliable catch information (A = Anchorage, S = Steaming, T = Trawling and E = Entry/Exit).

8.2.2 *Fitting of Model*

The HMM is fitted to the data, for each vessel independently, in an unsupervised learning sense to ascertain both the parameters $\Omega = \{\phi, P\}$ of the HMM and the transition matrix of the conditional probabilities $w_{tij} = \Pr[z_{t,i} = 1 \ \ \& \ \ z_{t+1,j} = 1 | y$. Taking the row sums of w_{tij} provides the posterior probabilities of the vessel belonging to each state at each time point i.e. $pr\{z_{t,j} = 1 | y\}$. These posterior probabilities obtained from the fitted model then provide a probabilistic measure of being in any given state. The HMM can be fitted via the Baum-Welch algorithm (Baum, Peterie *et al.* 1970; see Rabiner 1989) which can be described as a special case of the EM algorithm (Dempster, Laird *et al.* 1977). In essence the EM algorithm involves two steps, the E- and M-steps. In this application, the E-step estimates the transition matrix of probabilities w_{tij} given estimates of parameters Ω via a forward and backward algorithm. The M-step estimates the state parameters Ω given estimates of the transition probabilities w_{tij} . This circular algorithm is repeated refining the solution until some form of convergence is achieved.

In more detail, let S denote the number of states in the HMM (in our example S = 5), and we have observed vessel speed $y = \{y_1, ..., y_n\}$ over time 1 to n with time of day given by $d = \{d_1, ..., d_n\}$ and fishery given by $c = \{c_1, ..., c_n\}$ from a total of C fisheries. The algorithm is then given as:

E-Step

Forward Recursion The values $a_{i,t}^{(k)} \left(i = 1, ..., S; t = 1, ..., n-1 \right)$ are calculated on the $(k+1)^{\text{th}}$ iteration as follows;

Initialisation

$$a_{i,1}^{(k)} = P_{0i}^{(k)} f_i^{(k)} \left(\mathbf{y}_1, \theta_{ic_1} \right)$$

Induction

$$a_{i,t+1}^{(k)} = \left[\sum_{h=1}^{S} a_{ht}^{(k)} P_{hi}^{(k)} \left(d_t, c_t \right) \right] f_i^{(k)} \left(\mathbf{y}_{t+1}, \theta_{ic_{t+1}} \right)$$

Backward Recursion The values $b_{i,t}^{(k)} \left(i = 1, ..., S; t = n-1, ..., 1 \right)$ are calculated on the $(k+1)^{\text{th}}$ iteration as follows:

Initialisation

$$b_{i,n}^{(k)} = 1$$

Induction

$$b_{i,t}^{(k)} = \sum_{h=1}^{S} P_{hi}^{(k)}\left(d_t, c_t\right) f_i^{(k)}\left(\mathbf{y}_{t+1}, \theta_{ic_{t+1}}\right) b_{i,t+1}^{(k)}$$

Also within the E-step we estimate the posterior probabilities w_{tij} of moving from node i to node j at time point t,

$$w_{tij}^{(k)} = \frac{a_{i,t}^{(k)} \pi_{hi}^{(k)} f_i^{(k)}\left(\mathbf{y}_{t+1}, \theta_{ic_{t+1}}\right) b_{j,t+1}^{(k)}}{\sum_{h=1}^{g} \sum_{i=1}^{g} a_{h,j}^{(k)} \pi_{hi}^{(k)} f_i^{(k)}\left(\mathbf{y}_{t+1}, \theta_{ic_{t+1}}\right) b_{i,j+1}^{(k)}}$$

Let $w_{.jt}$ denote the probability of the vessel being in state j at time t, then

$$w_{.jt}^{(k)} = \sum_{i=1}^{g} w_{ijt}^{(k)}$$

Since the trawling state distribution is a mixture model, we must also estimate the posterior probabilities of each of the normal mixture components, i.e.

$$\tau_{ij}^{(k)} = \frac{\pi_j N\left(y_i; \mu_{2,j}, \sigma_{2,j}\right)}{\sum_{h=1}^{g} \pi_h N\left(y; \mu_{2,h}, \sigma_{2,h}\right)}$$

M-Step

The M-step consists of updating the parameter estimates of $\Omega = \{\phi, P\}$ based on the posterior probabilities estimated in the E-step. For state i and j,

$$P_{0i}^{(k+1)} = w_{i1}^{(k)},$$
$$P_{ij}^{(k+1)}\left(d, c\right) = \sum_{\forall t, d_t = d, c_t = c} w_{ijt}^{(k)} \Big/ \sum_{\forall t, d_t = d, c_t = c} w_{.jt}^{(k)}$$

where the summation over $\forall t, d_t = d, c_t = c$ denotes the sum over all t where $d_t = d$ and $c_t = c$. Next the state parameters are estimated, so for the steaming node/state where (c=1,…,C),

$$\mu_{5c}^{(k+1)} = \sum_{\forall t, d_t = d, c_t = c} w_{.it}^{(k)} y_t \Big/ \sum_{\forall t, d_t = d, c_t = c} w_{.it}^{(k)}$$

and

$$\sigma_{5c}^{(k+1)} = \sum_{\forall t, d_t = d, c_t = c} w_{.it}^{(k)} \left(y_t - \mu_5^{(k+1)}\right)^2 \Big/ \sum_{\forall t, d_t = d, c_t = c} w_{.it}^{(k)}$$

Notice that in this case we are not using all available information to estimate μ_{5c}, as the entry/exit 'data' also provide information on the parameter μ_{5c}. However, for simplicity and to avoid spurious model fits we choose to let the steaming node only estimate μ_{5c}.

In the case of the trawling node/state, for $(h = 1, \ldots g)$:

$$\pi_{hc}^{(k+1)} = \sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)} \tau_{ht}^{(k)} \Bigg/ \sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)}$$

$$\mu_{2,hc}^{(k+1)} = \sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)} \tau_{ht}^{(k)} y_t \Bigg/ \sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)} \tau_{ht}^{(k)}$$

and

$$\sigma_{2,hc}^{(k+1)} = \frac{\sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)} \tau_{ht}^{(k)} \left(y_t - \mu_5^{(k+1)} \right)^2}{\sum_{\forall t, d_t = d, c_t = c} w_{.2t}^{(k)} \tau_{ht}^{(k)}}$$

In the forward and backward steps, due to numerical issues when multiplying a large number of small numbers, all calculations were done on the log scale and long sequences of vessel inactivity were removed. Also, some restrictions were placed on the model parameters ϕ and some man-handling of posterior probabilities was included (for example, extremely high trawl speeds were assigned completely to the steaming node). The rationale was to ensure the model behaved and that no label-swapping occurred (i.e. when the model uses the steaming node to fit trawling and the trawling node to fit steaming).

8.3 Example

The approach described in this chapter has been applied to VMS data from Queensland’s otter trawl fisheries. VMS has been in operation in the fishery for six years, covering a fleet of up to approximately 500 trawl vessels. The fishery covers a range of 2000 km of coastline from the New South Wales border to the tip of Cape York. The fleet mainly targets the EKP (*Penaeus plebejus*), tiger/endeavour prawns (*Penaeus esculentus*, *P. semisulcatus*, and *P. monodon*/ *Metapenaeus endeavouri* and *M. ensis*), banana prawn (*Fenneropenaeus merguiensis*) and scallop (*Amusium japonicum* and *A. pleuronectes*). Figure 8.3 shows a histogram of calculated speed for a selected vessel, for which the trawling and steaming modes are clearly visible at approximately 1.3 m/s and 3.5 m/s. The non-Gaussian nature of the trawling speeds is also visible, illustrating the rationale for adopting a mixture to model trawling speeds.

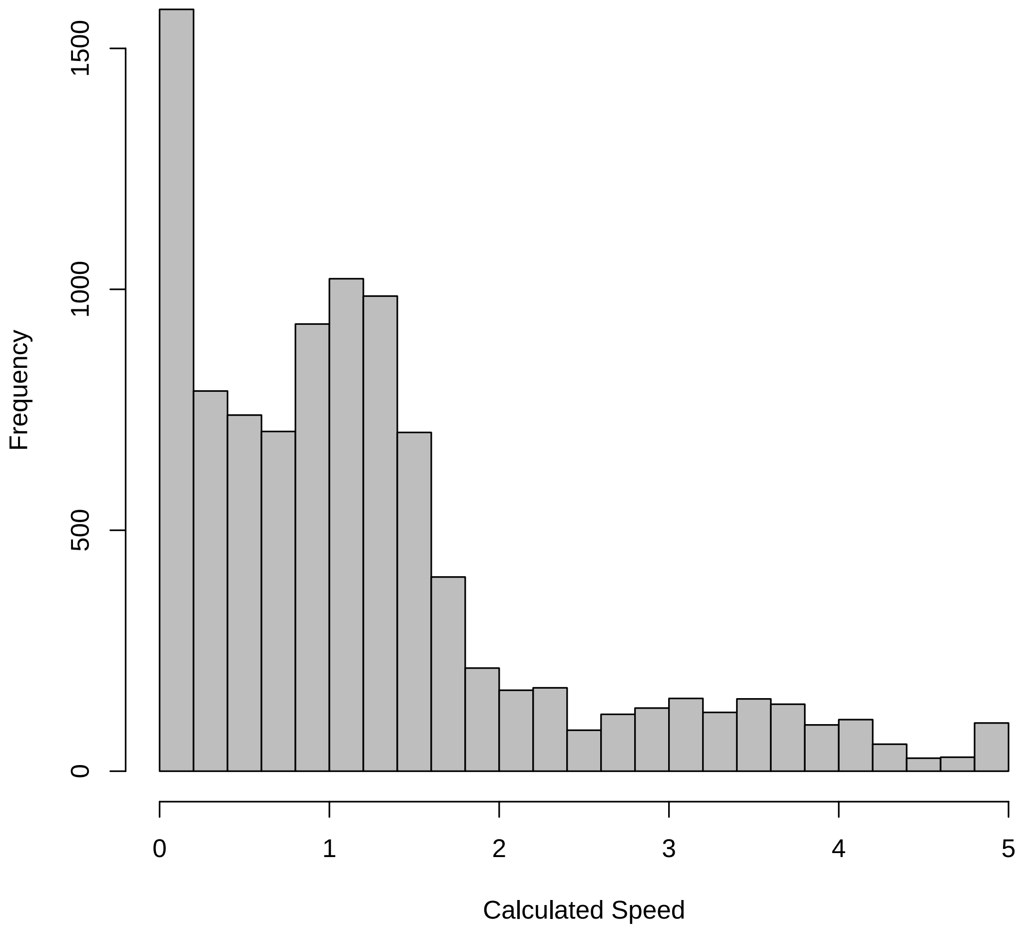


Figure 8.3 Histogram of calculated speed (m/s) for a selected vessel between 2000–2004.

Examples of VMS data from the EKP and scallop fisheries are given in Figure 8.4. A simple cut-off on the time of day and calculated speed was applied to provide a rough indication of trawling behaviour. It is obvious from these plots that the trawling behaviour can be quite different between fisheries. The allocation to fisheries was based on a simple maximum catch rule. In some fisheries using most caught species would definitely not be appropriate and a more eloquent approach would be required. However, in the examples discussed in this chapter, the spatial separation of fisheries is such that we tend to have reasonable spatial separation of catch.

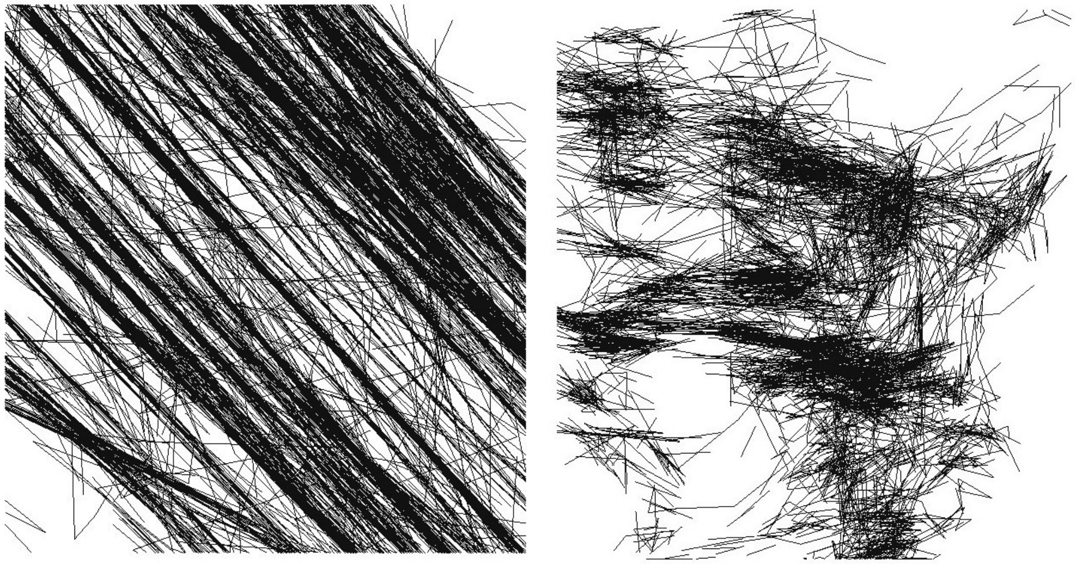


Figure 8.4 VMS data showing difference in trawl paths between two fisheries considered in this chapter: EKP (left) and scallop (right).

The method described in this chapter (implemented in FORTRAN 90) directly queries and updates the VMS oracle database to produce a probability of trawling for each poll. The raw VMS data were joined to logbook records of catch to establish the catch/effort of each fishery. It is possible that the logbook entries contain errors, in this case there may be valid trawling which does not match logbook entries. To avoid this problem VMS polls were allocated to the nearest logbook entry in time. The results where valid trawling appeared to be occurring in the absence of a logbook can possibly be used to validate or correct logbook data.

Due to the highly confidential nature of VMS data, spatial noise has been added, and a transformation has been applied to the results presented in this chapter, and some data in sensitive areas has been removed. The computation time was quite reasonable even for long time periods and in the order of minutes for a single vessel. The fitting procedure should have to be done only once and then the posterior probabilities stored and used from then on to determine vessel activity. As new data are made available either a new fit to the complete time series could be done, or if various conditions hold, the new data can be quickly assigned based on the existing HMM model parameters.

During development of this methodology the complete six years of data were available for the whole fleet. However, at the time of writing this chapter, only a single year of the data was available for the whole fleet, and for a small subset of vessels, data were available for four years. As such, the results for the example are split into two sections – for individual vessels and for the fleet.

8.3.1 *Individual vessel results*

The HMM was applied to selected vessels for which four years of data were available. The larger time period warranted and allowed the use of the multi-fishery version of the model. Figure 8.5 shows a representation of the resulting fit to a section of data. The polls have been allocated in an outright classification to states for clarity. The *x*-axis corresponds to time and the various symbols and letters indicate the fitted outright state. The *y*-axis shows the calculated speed. It can be seen that the model successfully determines the high speed polls as steaming, the near stationary polls as anchorage and the poll in the band between the two as trawling. Using a HMM has also successfully indicated several entry/exit polls. The temporal aspect of the model can clearly be seen in Figure 8.5. For example in Figure 8.5 (B), the poll at approximately the 115th hour has been assigned to the anchorage node. This is due to the influence of the time of day and the assignment of the two neighbouring polls. However, the calculated speed considered alone would, we would argue incorrectly, indicate trawling.

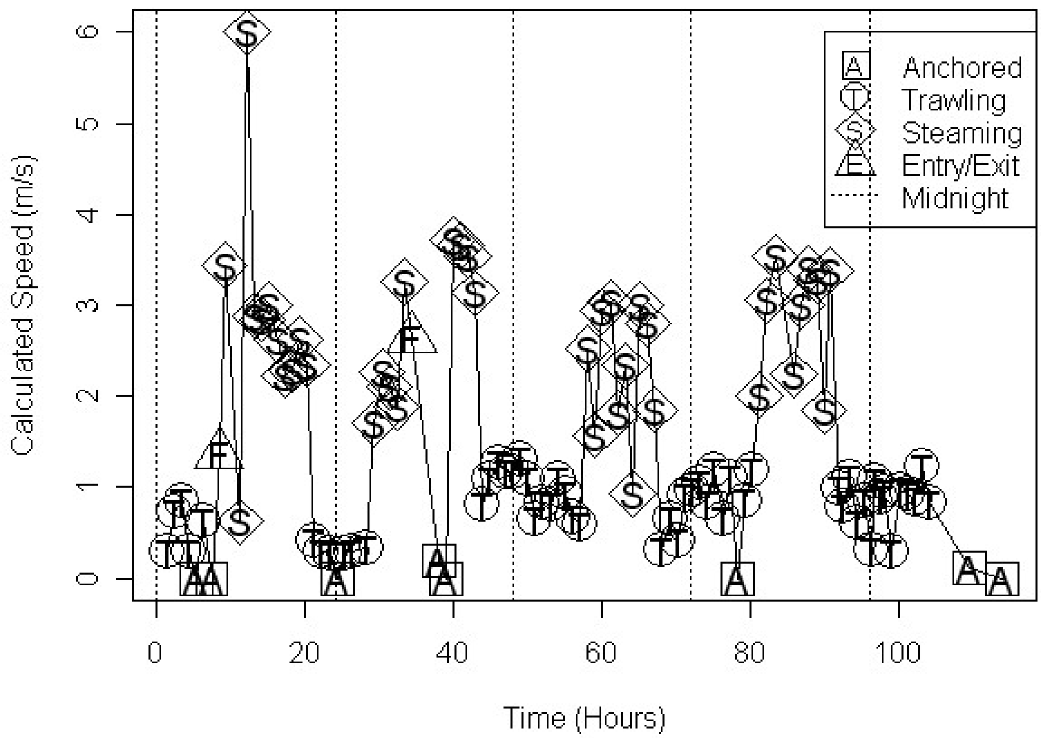
In practice it is possible that there may be a lack of historical data for some vessels, e.g. if a new vessel joins the fleet. In this case default model parameters Ω based on all available boats could be used. Another issue of data size may occur when a vessel switches to a different fishery for a very short time only. The easiest solution to this problem is to pool the relevant data together with another fishery for that vessel. For simplicity in this example, vessels with very small sample sizes were simply removed from the analysis.

If the subset of vessels examined here is indicative of the whole fleet then including the fishery effect may not have been as important as first thought. This is due to the vessels in our subset generally not changing between drastically different fisheries, e.g. the two extremes shown in Figure 8.4. So for the subset of vessels examined, including a fishery component did not substantially change the final result.

Figure 8.6 and Figure 8.7 present the results for a selected vessel. During the four years the vessel targeted two fisheries, scallop and EKP. Figure 8.6 corresponds to when the vessel was deemed to be fishing in the scallop fishery and Figure 8.7 similarly for the EKP fishery. In each figure the state distributions are denoted as a mixture and are overlaid on a histogram of calculated speed corresponding to the respective fishery.

It should be emphasised that the HMM is not based on calculated speed alone, but also includes time of day and the temporal correlated aspects of the data. These plots of raw calculated speed and state distributions do not convey these important factors, but are still useful to visualise the resulting fits.

(A)



(B)

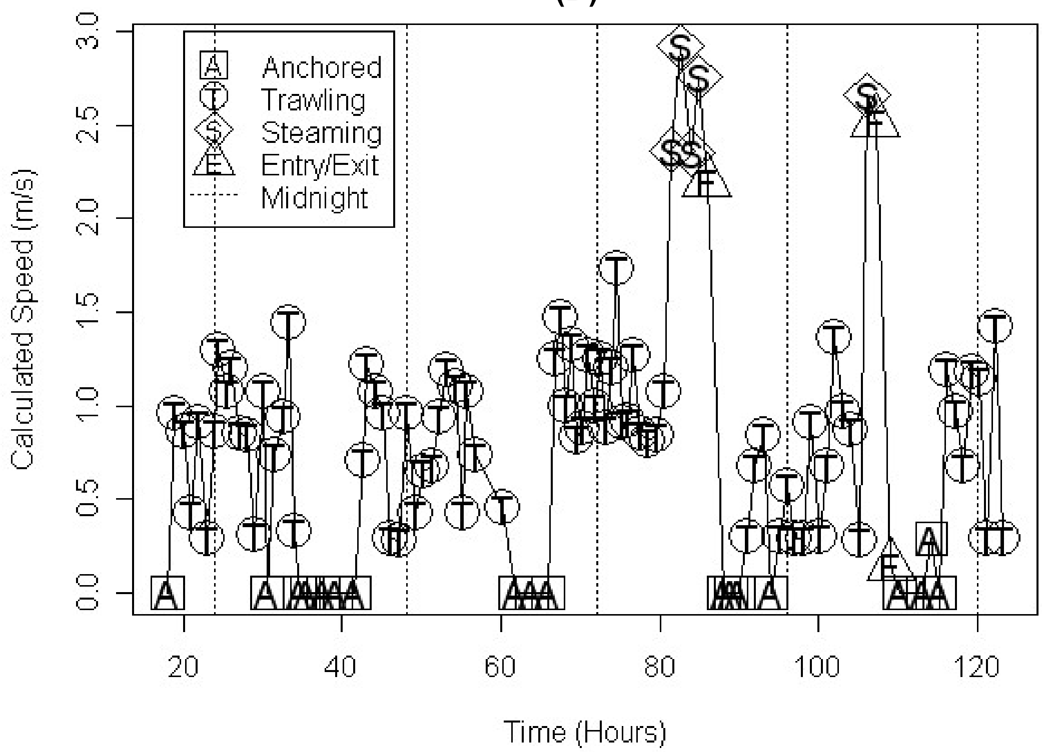


Figure 8.5 Example of a typical fit to a section of time for a single vessel A (top) and a vessel B (bottom).

The difference in trawl behaviour can be clearly seen in the histograms of calculated speed. In the scallop fishery there are many more lower trawl speeds due to shorter, less straight trawling behaviour. In contrast, for the EKP fishery there is a much more distinct mode corresponding to trawling due to longer, straight trawling behaviour. This can be seen in the plots of trawl tracks in Figure 8.8. For this vessel the actual speed corresponding to the mode of trawl speed does not seem to be that different between the fisheries.

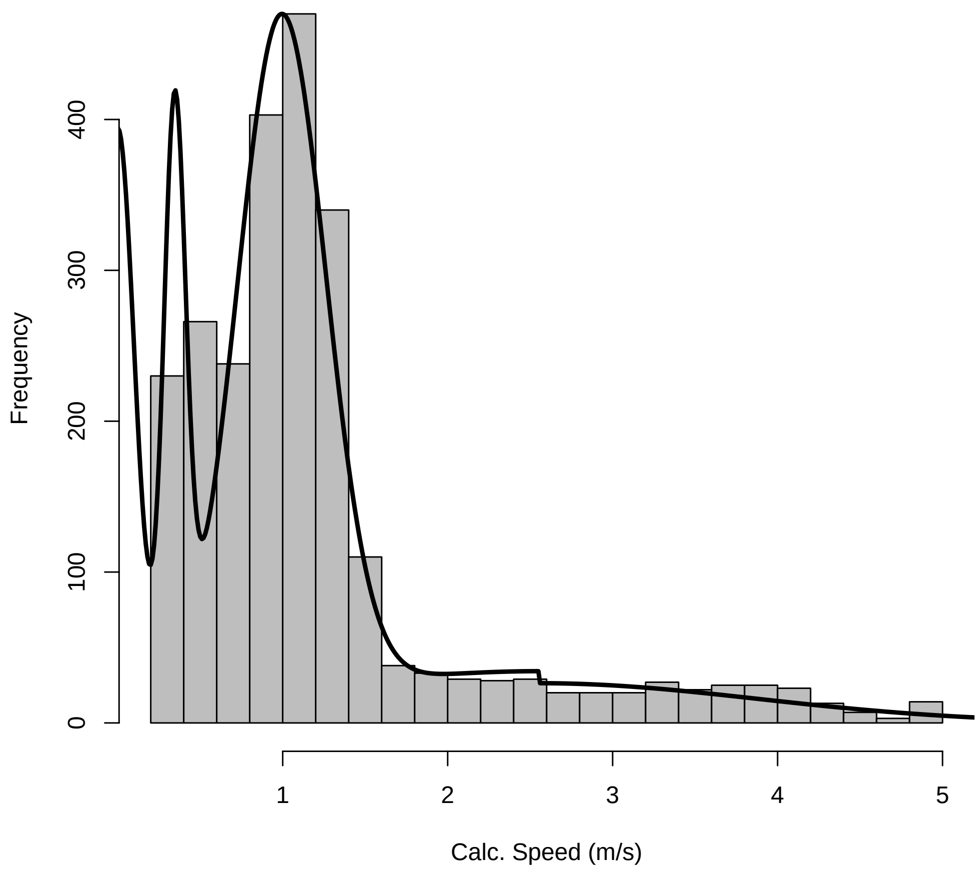


Figure 8.6 Example of the fit for a single vessel of calculated speed corresponding to when the vessel is deemed to be in the scallop fishery.

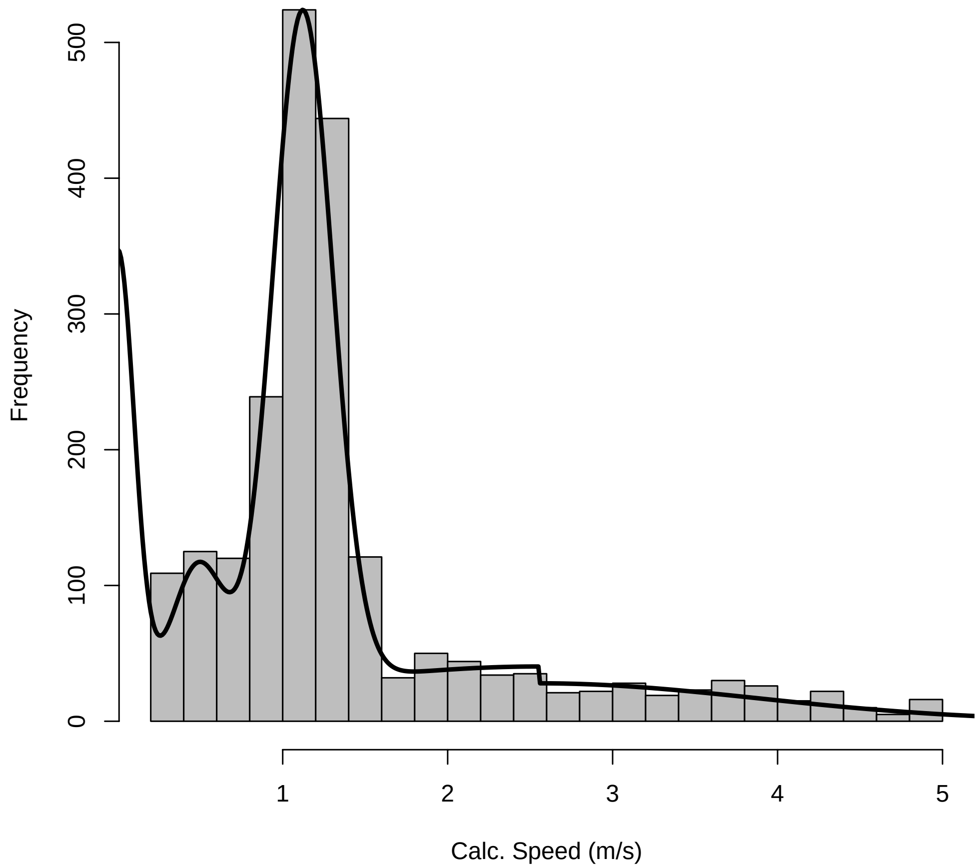


Figure 8.7 Example of the fit for a single vessel of calculated speed corresponding to when the vessel is deemed to be in the EKP fishery.

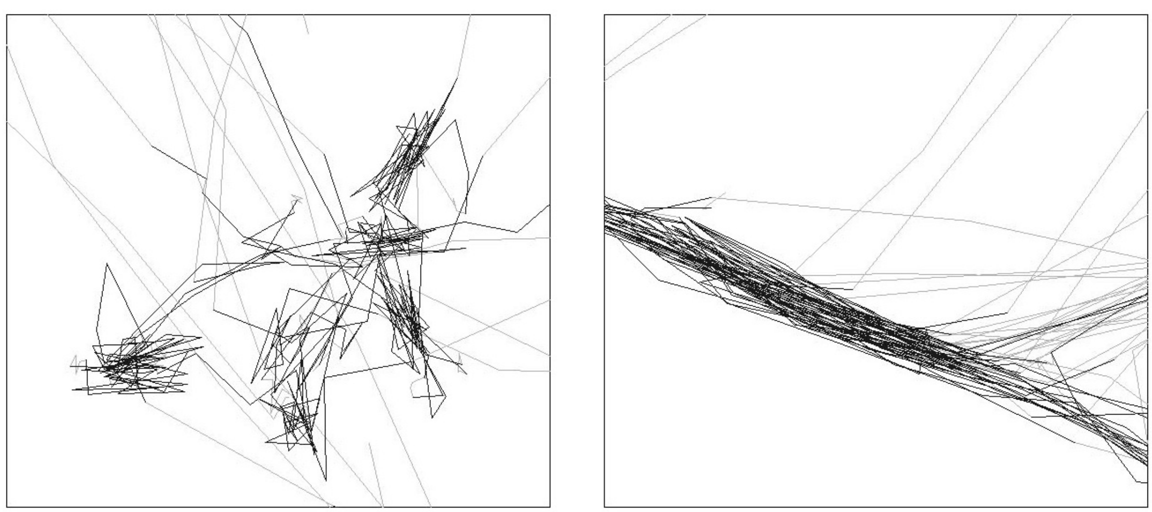


Figure 8.8 Plot of actual tracks for the selected vessel allocated to fishery. The left plot indicates tracks in the scallop fishery and the right the EKP fishery.

In Figure 8.6 there appears to be an issue with the fit of the mixture model related to trawling. This can be seen as the high mode at lower trawl speeds that does not match the underlying histogram. Although the overall answer with regard to classification of trawling was not greatly affected, it does illustrate that some caution should be taken with any automated approach – in particular, that spurious models are not being found by the fitting process. To avoid spurious solutions the use of a range of different starting values may be required. As can be seen in Figure 8.9, fitting a normal mixture assuming equal variances does solve this specific issue and provide a more natural fit.

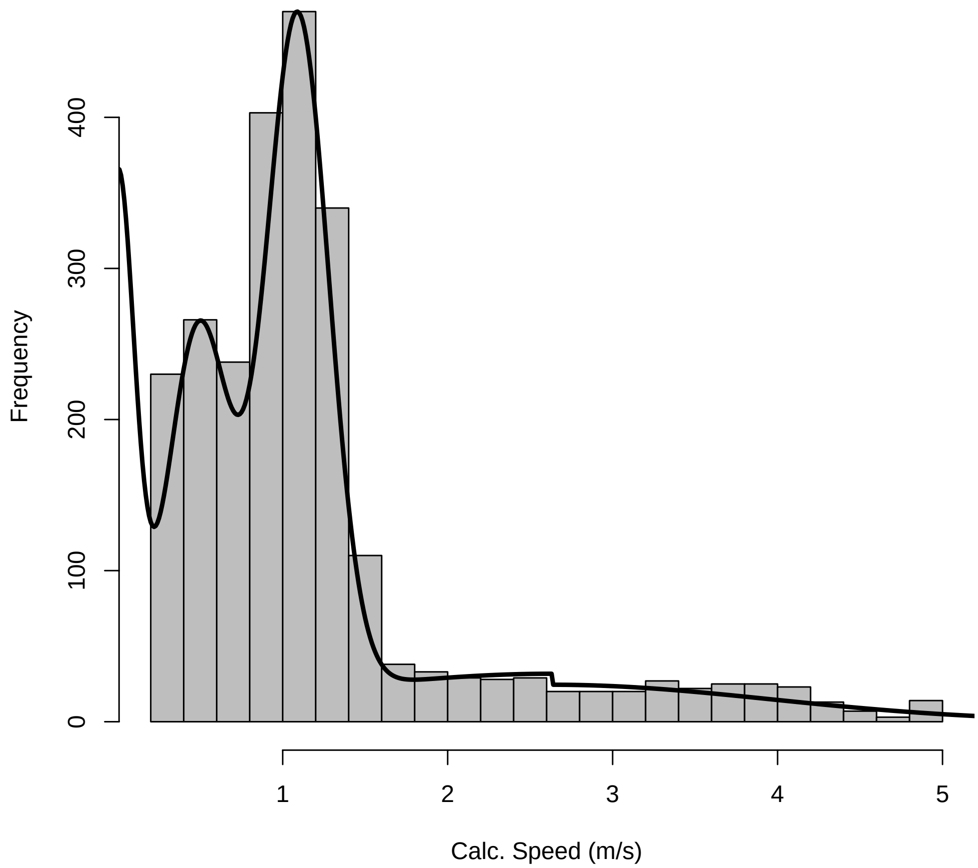


Figure 8.9 Example of the fit for a single vessel of calculated speed corresponding to when the vessel is deemed to be in the scallop fishery constraining the trawling mixture model to have equal variances.

8.3.2 *Fleet-wide results*

The HMM was applied to the whole fleet, based on a single year of the available data. As long as numerical difficulties can be avoided and there is no unmodelled temporal change in vessel behaviour over the full time period, fitting the models to the whole data set should, if anything, improve the results. Due to the relatively short time period, most boats only targeted a single fishery, and in the case of the vessels that did change fisheries, there was insufficient data to fit the fishery model. So, for this whole fleet example, we fitted the single fishery version of the model.

The probabilistic allocation to the trawling state can be used to produce spatial and temporal maps of fishing intensity for the whole fleet. Figure 8.10 shows an example of a section of estimated effort intensity for a single calendar year. The Queensland fisheries in question use a logbook system that has historically required precision at the degree level and more recently to one-tenth of a degree. For some indication of the relative increase in spatial resolution, Figure 8.10 is approximately ½×¾ of a degree. So the use of VMS data in this fashion provides a substantial increase in spatial resolution over logbooks alone. However, equally as important is that logbook entries correspond to a single point representing a full 24-hour period of fishing. On the other hand, the VMS system provides location information on an hourly basis. So using VMS data substantially increases temporal resolution, which also contributes to increasing spatial precision.

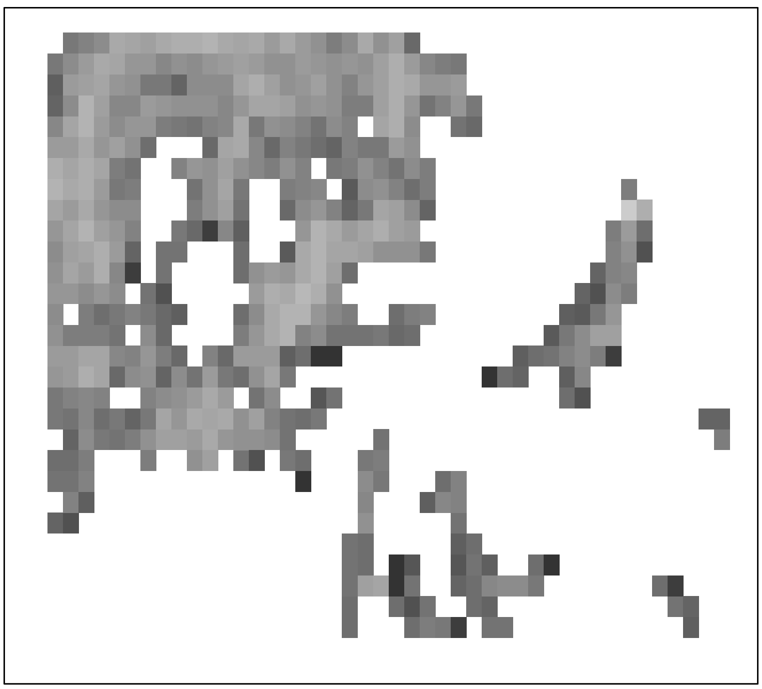


Figure 8.10 Example of effort density map produced for the EKP fishery.

Figure 8.11 and Figure 8.12 show selected regions of data after the HMM has been applied. The tracks where the trawling probability is greater than 0.5 are denoted by black lines, and the cases where trawling probability is less than 0.5 by grey lines. It would seem, on the whole, that the classifications are reasonable with distinct steaming and trawling visible.

Closer examination of the allocations does show some residual misclassified effort on the entrance and exit of an anchorage (see Figure 8.13). The actual percentage of data misclassified around the anchorage is very small. However, further work may be required to completely correct these errors, as detailed in the discussion section.

Figure 8.11 Plot of a region of VMS data for a single year with HMM applied. Transects with trawl probability < 0.5 denoted by grey lines and trawl probability ≥ 0.5 denoted by black lines.

Figure 8.12 Plot of a region of VMS data for a single year with HMM applied. Transects with trawl probability < 0.5 denoted by grey lines and trawl probability ≥ 0.5 denoted by black lines.

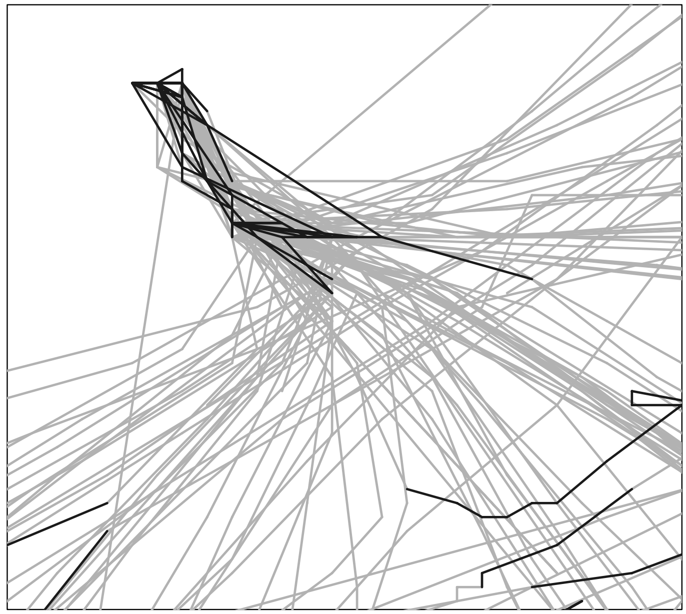


Figure 8.13 Plot of VMS data for a year in a particular region with HMM applied. Transects with trawl probability < 0.5 denoted by grey lines and trawl probability ≥ 0.5 denoted by black.

8.4 Validation/Error Quantification

8.4.1 *Relative performance of HMM*

In this section we examine a measure of error to establish the performance of the method. The only related work that could be found in the literature was part of the work to establish if VMS could be used to estimate depletion (Deng, Dichmont *et al.* 2005). In this work, a simulation study was presented to examine the effect on statistical power of polling frequency. In this chapter, we present a comparison between using a HMM and a simple speed cut-off rule. This was done using fishers logbooks that recorded individual trawl shots/tows as validation to produce an error rate.

A small number of vessels in the Queensland trawl fleet record their logbook catches on an individual trawl basis rather than the usual nightly summary catch record. These data provide a start time, time trawled and an approximate trawl start position. Matching these ‘true’ data with the VMS data provides a test data set consisting of raw VMS polls and a corresponding indication of the true vessel activity. From these data we can obtain estimates of error rates for our methods. It should be noted from the outset that the logbook data may have varying degrees of accuracy and in some cases may not provide a true reflection of actual fishing at a given VMS poll. For example if a fisher misreports a single shot this would give approximately 2 to 4 data points where 100% error may occur even if the method is working perfectly. However, as a relative measure of error to compare approaches these data should be acceptable.

Each trawl logbook entry was mapped to the VMS data corresponding to the same time period. This resulted in a measure of the proportion of the polling time that was spent trawling for each poll. We filtered the data for obvious fisher error (e.g. overlapping entries) and trawls of unrealistic extreme length and produced a test data set of 17,291 VMS polls from between 2000 to 2004 consisting of 11 vessels. Approximately 30% of the records corresponded to scallop, 46% to EKP, 1% to banana prawn, and 12% to tiger/endeavour prawns catches in terms of largest catch. For ease of calculation, a single fishery was assumed in the HMM model for each vessel.

Table 8.1 shows a comparison between the error rates of HMM and of a simple outright classification based on calculated speed. The error reported is calculated as Time Misclassified (secs)/Total Time (secs). The HMM approach shows a consistent decrease in error rate for all vessels.

Table 8.1 Comparison of error rates based on shot-by-shot logbook records.

		Error (%)	
Vessel	No. of Polls	Filter Method	HMM Method
1	4463	39.27	37.82
2	159	47.23	45.38
3	49	28.35	27.79
4	319	44.17	37.54
5	4463	39.27	37.82
6	190	36.62	31.64
7	336	58.45	63.28
8	102	30.08	22.58
9	571	46.47	41.06
10	376	42.42	40.57
11	6263	40.91	39.55
Total	17291	40.59	38.99

The methods we have developed at present assume that the trawl path between polls is a straight line. However, since there is a significant time interval between polls there is some spatial uncertainty of vessel positions between polls. The amount of uncertainty will depend on fishery, for example, due to the straightness of trawl paths or length of trawl tows.

8.4.2 *Examination of trawl path assumption*

An initial investigation into the validity of the assumption that vessels travel in a straight line between polls was conducted. To accomplish this we obtained available high-frequency VMS (polling less than 5 minutes) and fishers' GPS logger data. These data was used to produce empirical distributions of vessel location between artificially sampled polls.

The spatial distribution between two consecutive polls depends on spatial distance between polls. In particular, polls that are close spatially will have more spatial uncertainty than polls that are further apart, assuming a reasonably constant speed. Only a small study has been completed but upon a cursory examination the uncertainty is not as high as expected, especially between polls that are well separated. Figure 8.14 shows one such empirical distribution. The frequency has been logged so the actual distribution is much less spatially dispersed than in this figure. The spurious long lines are due to misclassified steaming being included. From this study it would seem that with hourly polling the straight line is a reasonable approximation. However, it would be reasonably simple to instead use the empirical distributions arising from high-frequency data studies like this to better convey the spatial uncertainty.

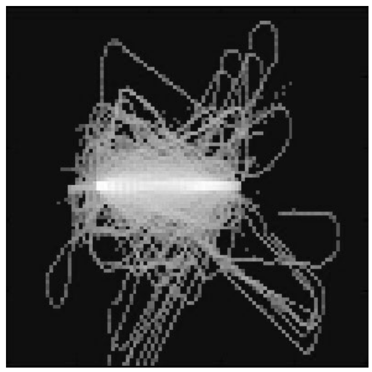


Figure 8.14 Empirical distribution of vessel location between polls at close to a fixed distance apart (the darker colours indicate low probability and the lighter colours indicate high probability.)

8.5 Discussion

Utilising a Hidden Markov Model provides an automated method to determine vessel activity at an individual vessel level. Due to its very nature it models the innate temporal correlation of the data very well. We have previously developed more complicated approaches combining several methods or steps, plus some ad hoc adjustments. However, the beauty of the HMM approach is that most of the issues with the data are addressed within a single simple statistical framework.

Each vessel in each fishery has its own HMM to reflect individual peculiarities of vessels and skippers. Since the fitting of the HMM is reasonably automated it is feasible to use the approach on large fleets. Obviously, compared to using a fixed pre-defined speed cut-off, more care must be taken with an automated fitting procedure. In particular, care must be taken to check for spurious solutions, or that data errors are not influencing the model fit. In general it would be advisable to assess the individual vessel fits to confirm that reasonable fits are being obtained.

From the example presented, it can be seen on the whole that the HMM model works very well. The method produced quite intuitively reasonable estimates of fishing effort. However, there are some aspects which could still be improved. In particular, when entering an anchorage there are still some residual polls misclassified as trawling when the vessel's navigation from the anchorage spans two polling intervals (i.e. if the anchorage is located inland on a river, or if shoals or islands must be navigated around). Using the actual data there is probably no straightforward automatic way to completely discern this behaviour from fishing. The simplest approach may be through post-analysis, spatially determining these anchorage pathways via visual inspection of maps. Alternatively, a more automated procedure would be to spatially identify high concentrations of entry/exit classified polls and simply re-classify polls that are in the same location.

Spatial effort can be mapped at varying spatial or temporal scales. If spatial effort is examined over a longer time period some form of standardisation of effort would obviously be required (Bishop, Venables *et al.* 2004; Maunder and Punt 2004; O'Neill, Courtney *et al.* 2003). Care should also be taken not to overestimate the spatial accuracy of the resulting effort distributions. Since the VMS is accurate to 50 m it is tempting to examine the data at an extremely fine spatial scale. However, the spatial uncertainty of vessel positions between polls should be considered, as seen in the previous section. To better portray this uncertainty a natural progression from the assumption that vessels travel in a straight line is to use high-frequency data, such as GPS logger data, to produce empirical distributions of vessel location between polls depending on spatial distance between polls.

A comparison between using a HMM and a simple speed cut-off rule was conducted on a sub-sample of the data where logbook data on individual trawls were available for validation. Although the overall improvement in error rate by using an HMM is small, across the whole fleet over time this would correspond to a large amount of misclassified effort. It should be again noted, that the definition of the logbook entries reflecting 'true' vessel activity may be problematic. Hence a portion of this misclassification error may be impossible to avoid. The most pathological example of this is two states with a very small number in one. In this case a very low misclassification error can be obtained with the classification rule of assigning all observations to the state with the larger number of observations.

For the assumed state distributions, there is possibly a better choice for the distribution arising from the entry and exit states than uniform distributions. The entry/exit state corresponds to a vessel that spends a proportion of the polling period stationary and the remainder at steaming speed. Hence, a distribution that better reflects the distribution of the sum of a uniform distributed random weighting of the steaming and anchoraging states' speed distributions would be more appropriate. This is not to be confused with a mixture of these two distributions. The inappropriateness of the uniform can be seen by the non-smooth nature of the overall distribution at the upper limit of the entry/exit uniform in Figure 8.6, Figure 8.7 and Figure 8.9.

The HMM approach lends itself easily to a more complex model if required. For example the model could be adapted to estimate targeting as well as vessel activity. In this example it was assumed that the most caught species and the spatial location on any given trawl indicated the target species. This is true to a large extent, however it's not perfect and obviously the targeting behaviour of fishers will be temporally correlated. This could easily allow additional covariates to be included that help to predict targeting behaviour, possibly in a design similar to Figure 8.2. Other information or data that helps to predict vessel activity (e.g. weather, ocean currents, vessel direction) can be very easily incorporated into the HMM model as well.

In some VMS implementations, polling frequency may vary. For example, when a vessel nears a closure or protected area some systems increase polling frequency. In this case the model would have to be modified to accommodate different time steps. One approach could be to re-frame the model as a continuous time Markov model.

In summary we have found that using a HMM approach successfully addresses many of the issues in VMS data, such as temporal correlation and misclassification of entry/exit to anchorage as trawling. In the example provided it was found HMM gave improved results compared to a speed-based cut-off rule. Overall HMMs provide a powerful elegant tool to better extract the wealth of information available in VMS data.

9 USING MAXIMUM ENTROPY METHODS TO CALCULATE AN INDEX OF ABUNDANCE FOR PRAWN STOCKS, BASED ON VMS-DERIVED CATCH AND EFFORT DATA

Norm Good and David Peel

9.1 Introduction

Use of VMS for compliance in global fisheries is increasing at a considerable rate. The vessel position information provided by VMS can be a valuable source of data for fisheries management. By combining information on fishing vessel location with logbook catch records, maps recording fish catch, fishing effort and CPUE can be made at very fine spatial scales. These resulting maps can then be used in fisheries resource assessments to identify local fish aggregations and how they change over space and time. The finer scale CPUE data can also be used for more accurate fish stock assessments because it removes the effects of hyperstability inherent in using CPUE based on large spatial grids (Campbell 2004; Campbell, Tuck *et al.* 1996). However, the problem of scale remains. For example, a vessel's position may be recorded hourly by a VMS but logbook entries of catch may be done daily. The resulting maps based on catch records thus may be a blurred representation of the actual resource. When catch data are recorded at a larger spatial scale than the scale desired, the use of geostatistical techniques to calculate stock density may be more applicable than using average CPUE. However, many of these techniques do not perform well when a highly aggregated resource is mapped (Maravelias, Reid *et al.* 1996). These methods tend to treat some parts of highly aggregated concentrations as outliers because of the difficulty in modelling them (Rivoirard, Simmonds *et al.* 2000).

Methods such as kriging rely on linear functions (i.e. variogram models) to describe the relationship between neighbouring points. In turn, these assume a gradual change from one point to another. Use of linear functions is acceptable if there is sufficient fishery-independent sampling coverage (i.e. regular sampling taken over a study area and based on a grid system). For schooling species, sampling of stock density is usually carried out using sonar, echo sounders, or other cost-effective remote sampling tools. However, such tools are not feasible for sampling demersal species such as prawns (Penaeidae), and as well, fishery-independent sampling using trawling is expensive. Estimates of stock density or stock size, therefore, are usually obtained from commercial logbook catch and effort data.

Using commercial data to conduct geostatistical analysis can be problematic. Murray (1996) provided a classic fisheries example of a highly skewed resource, such as the Antarctic krill (*Euphausia superba*), where geostatistical techniques failed on commercial data.

Whilst position information is available for individual tows, the majority of fishers record daily catch. Such is the case for the majority of Queensland EKP logbook catch records, in which locality information is based on a six-minute (coordinate) grid, and used to record the position of the largest catch. Trawling generally takes place between dusk and dawn with individual trawl shots lasting up to four hours at speeds of up to four nm per hour. Consequently, the location of catches at a scale finer than the distance a trawler can travel in a day cannot be precisely obtained.

To map catches and associated indices at finer spatial scales, the distribution of the catch over a night's fishing needs to be estimated. This can be determined if a number of individual tows and associated catches have crisscrossed each other within a relatively short period of time, and enough of them are contained within a relatively small region. However, a technique (such as probability modelling) for estimating resource intensity that does not rely on linear modelling would be more useful.

The maximum entropy principle (MaxEnt) has been used in a number of fisheries applications. Vignaux *et al*. (1998) applied it to estimate fish density in the New Zealand ‘hoki’ (*Macruronus novaezealandiae*) spawning fishery, and Brierley *et al*. (2003) successfully applied it on acoustic survey data to map the density and biomass of Antarctic krill. To obtain a relatively clear picture of the stock distribution, Vignaux *et al*. (1998) used relatively large spatial grids of 8 × 8 nm as it was difficult to identify temporal trends and local patterns using grids of smaller sizes (Lizamore 1995). In our study we use the maximum entropy method to estimate the resource distribution of EKP stocks in Queensland at scales down to 1 × 1 nm, by incorporating a spatial correlation function within the MaxEnt formulation.

9.2 Methods

9.2.1 *Maximum Entropy Theory*

Maximum entropy analysis originated in the 18th century firstly through the works of Bernoulli and expanded upon by Laplace. Jaynes (1980) provides a detailed account of these developments. Burg (1975) was the first to use it for data analysis when applying it to spectral analysis. Maximum entropy is a technique based on probability theory for reconstructing distributions, and is particularly suited to handling noisy and sparse data consistently (Gull and Skilling 1999). The basic theory behind maximum entropy is that lacking information about a certain quantity, we assign that quantity several possible probability distributions, each of which satisfies constraints defined by our prior knowledge. We then select the distribution that maximises the entropy, S.

The technique was first applied to the reconstruction of fuzzy images by Gull and Daniell (1978), by reducing background noise whilst sharpening the major parts of an image. A well-known example of this application is the reconstruction of the images captured from the flawed Hubble space telescope (Llacer and Núñez 1990). However, applications more relevant to fisheries' problems, such as that described above where we have crisscrossing tows, include areas such as X-ray tomography. In this process, X-rays are passed through the body from a number of directions and the absorption intensity of each X-ray line is measured. The MaxEnt method uses the information from a large number of scans to map areas of high absorption.

Whilst MaxEnt is used to assign a prior distribution to initially model fish density, Bayesian techniques are used to update the posterior distribution. Bayes's Theorem in a maximum entropy context can be defined as:

$$P(hypothesis \mid data) = P(hypothesis)P(data \mid hypothesis) / P(data)$$

where $P(hypothesis)$ is the prior probability assigned to the distribution of fish density using maximum entropy; $P(data \mid hypothesis)$ is the probability of obtaining the data given a certain prior density (also known as the experimental likelihood, which often is in a Gaussian form); and *P(data)* is the probability of the data (also known as the evidence), which is used to test different hypotheses. The posterior probability $P(hypothesis|data)$ is the updated probability of the fish density. The main aim is to modulate the prior data so that the objective function (i.e. the posterior likelihood) is maximised using non-linear optimisation search algorithms.

Model formulation starts by dividing an image into n blocks. The image is represented by the density of fish h_j in each block j. Information is provided by the p tows that run over the image and associated catch D_i of the i^{th} tow. This is demonstrated below,

$$D_i = \sum_{j=1}^{n} R_{ij} h_j + \varepsilon_i$$

Equation 9.1

where R_{ij} is the response matrix that enables movement from image space to data space (Lizamore 1995) and is the distance in nm that tow i passes in block j in which there were h_j kg per nm of fish. ε_i is a random error distributed as $N(0,\sigma^2)$. This error is assumed to include a variety of unmeasurable effects such as variable catchability and observational errors.

To find h_j we can use D_i to find the most plausible values of h_j consistent with D_i using a Bayesian method, where the posterior probability of a particular set of h_j values (known as the candidate image) is based on both the prior probability — the plausibility of the image based on what we know about the data generally — and the likelihood of obtaining the data if the trial image was in fact the true image.

The MaxEnt method assigns a prior probability of a candidate image as

$$P(\mathbf{h}') \propto \exp(\alpha S(\mathbf{h}'))$$

Equation 9.2

where α is a regularising constant to be estimated, and $S(\mathbf{h}')$ is the information entropy:

$$S(\mathbf{h}') = \sum_i \left[h_i' - m_i - h_i' \log\left(\frac{h_i'}{m_i}\right)\right]$$

Equation 9.3

and m_i is a vector of the initial estimate of fish density, assumed uniform in this study.

The prior probability is modulated using a likelihood which, assuming normal errors, takes the form

$$P(\mathbf{D} \mid \mathbf{h}') = Z^{-1} \exp\left(-\chi^2/2\right),$$

Equation 9.4

where

$$Z = \prod_i \left(2\pi\sigma_i^2\right)^{1/2},$$

Equation 9.5

and

$$\chi^2(h) = \sum_i \left(\frac{\left(H_i' - D_i\right)^2}{\sigma^2}\right),$$

Equation 9.6

the chi-square statistic between the observed and predicted data, and,

$$H_i' = \sum_{j=1}^{n} R_{ij} h_j',$$

Equation 9.7

are the predicted values given the trial image, and σ^2 is an estimate of the variance.

Using Bayes's Theorem, the posterior probability of the trail image $\mathbf{h'}$ is

$$P(\mathbf{h}' \mid \mathbf{D}) \propto \alpha S(\mathbf{h}', \mathbf{m}) - \chi^2(\mathbf{h}')/2$$

Equation 9.8

To maximise the Bayesian posterior, we start with a trail image $\mathbf{h'}$ and move towards $\mathbf{h'}$ which maximises Equation 9.8, starting from the global maximum of entropy at $\mathbf{h} = \mathbf{m}$.

The software package MemSys5 was used to estimate values for α and σ^2 (see Gull and Skilling (1999) for methods).

The evidence term,

$$P(\mathbf{D}) = \sum_i P(\mathbf{D}, \mathbf{H})$$

Equation 9.9

is evaluated in MemSys5 using a conjugate-gradient method to test alternative hypotheses.

Fundamental to maximum entropy is that entropy maximisation should not itself introduce correlations between individual cells. However, in many cases it is known *a priori* that an image or school contains correlations (Gull and Skilling 1999). This knowledge can be incorporated into MaxEnt by deriving a quantity of interest (such as a school of prawns) as a blurred reconstruction of the underlying 'hidden variables', designated as f. Blurring is accomplished with an operator of the form

$$f_j = \sum_{i=1}^{L} C_{ji} h_i \quad (j = 1, 2, ..., M)$$

Equation 9.10

This operator is known as the intrinsic correlation function (ICF), where C_{ji} is a symmetric M × M matrix describing the blur. MemSys5 uses a cubic-spline approximation of a Gaussian point-spread function (Bontekoe, Koper *et al.* 1994), where the correlation between neighbouring grids for a given ICF width (ICFWIDTH) is

$$W_{width \geq 2} = \sqrt{\frac{ICFWIDTH^2 - 1}{3}}$$

Equation 9.11

The entropic prior h by itself is uncorrelated, as all correlation is assigned to the ICF. Depending on the situation, h and f can be reconstructed on different scales, and belong to two different spaces – hidden for h and visible for f – where L is the number of cells in the hidden space and M the number of cells in the visible space (Gull and Skilling 1999). Multiple scales can also be included to account for smooth (e.g. similar prawn densities in adjacent recruitment grounds) and sharp (e.g. prawn congregations along specific depth contours) reconstructions, though such features were not investigated in this chapter. MemSys5 allows you to test various ICF blurring widths against the evidence term to select the most probable model (see Brierley, Gull *et al.* 2003).

9.2.2 *Catch and effort data*

Figure 9.1 shows the extent of the EKP stock in Queensland. There are two separate fisheries: a shallow water fishery in areas adjacent to the coast and islands in depths less than 90 m; and a deepwater fishery in areas with depths greater than 90 m. The shallow-water fishing season occurs mainly during the summer months (November to February) and catches are predominantly of new recruits (carapace length greater than 26 mm) from estuarine nursery grounds. The deepwater fishery has its largest catches from May to August and mainly represents a relatively mobile stock of spawning individuals that migrate from south to north along the 90 m isopleth to Swain Reef (Courtney, Masel *et al.* 1995).

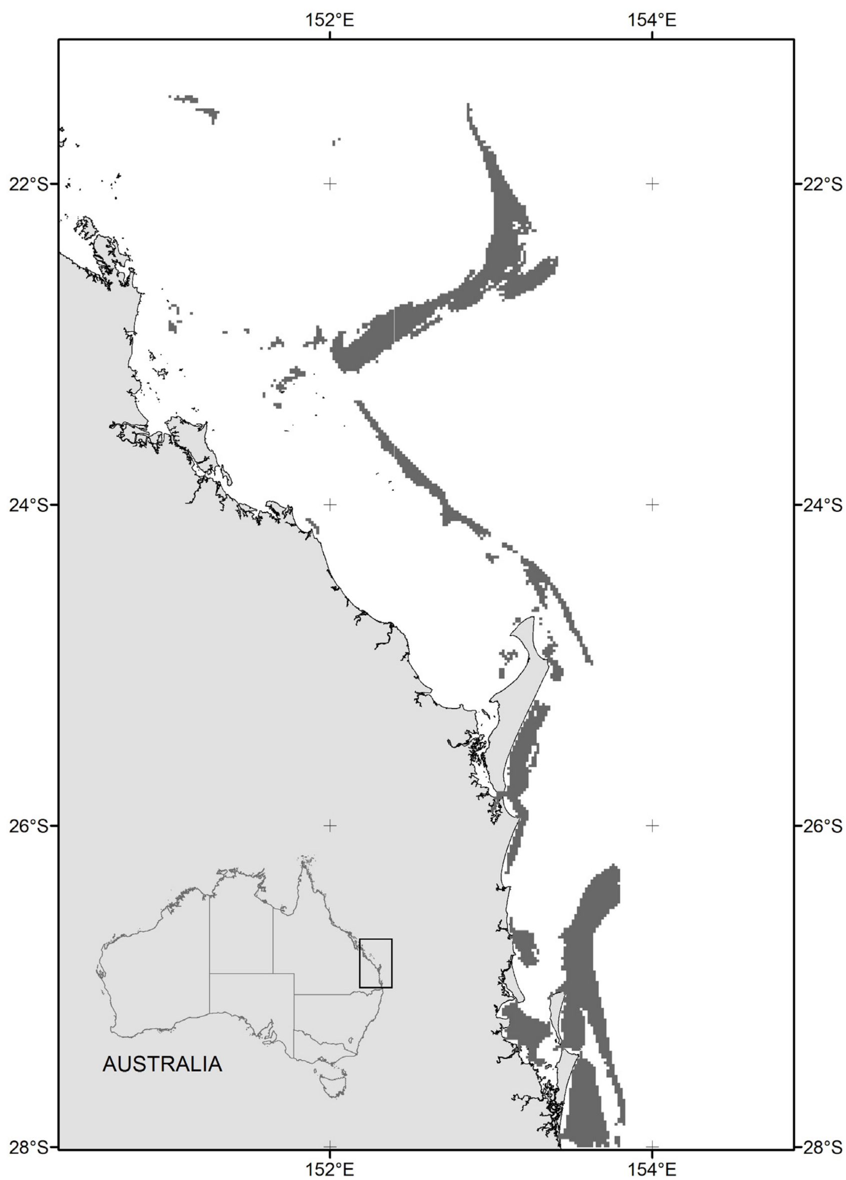


Figure 9.1 The extent of the EKP (shown as dark-shaded one nm square grids) from the border of New South Wales (29^{0}S) to Swain Reef (21^{0}S).

Daily catch records of EKP obtained from fishers' logbooks were matched to position information obtained from the VMS database, based on vessel number and date. The resulting VMS/Catch database consisted of a number of hourly position reports for each daily logbook catch record. These position reports were then joined to create individual hourly tow lines. To remove tows which were deemed not to be trawling, a number of decision rules were applied to filter the data (see Chapter 7). The data were then summarised by converting into individual grid cells of varying size. Average nightly trawling times after applying the filter ranged from 12.8 hours in summer months to 13.7 hours in winter months.

9.2.3 *Standardising catch data*

To estimate prawn density over the Queensland EKP stock area using the catch and effort data we needed to ensure that the catches from individual vessels were comparable. We used a standard approach to model catch using a generalised linear model (GLM) which incorporates vessel characteristics, seasonal effects, gross area, and effort, as linear predictors:

$$\log D_i = \sum_j \beta_{i,j} + \varepsilon_i$$

Equation 9.12

where β_{ij} is variable j having value i and ε_i is the normal and independently distributed [NID($0,\sigma^2$)] residual that may be explained partly by variation in prawn density at a local scale. Variables used in the model were: effort (hours trawled per night), year, 30 × 30-minute spatial grid, month, month × grid interaction, lunar phase, seven-day adjusted lunar phase (lagged phase to identify a waxing or waning moon based on lunar phase), vessel code, and scallop catch. To account for zero catches, a constant equivalent to 10% of the overall mean catch was added to all catches, including scallop (*Amusium ballotti*) catch (Campbell *et al*. 1996).

A stepwise GLM procedure was implemented in the GenStat 5 statistical package, Lawes Agricultural Trust, Rothamsted Experimental Station, UK (GENSTAT 5 Committee 1993) to select factors for standardising catch. All factors were found to be significant at the 5% level, explaining 69.1% of the variation in the catch. The resulting standardised back-transformed catches were then used in MemSys5 to investigate fine-scale structure in prawn density.

9.2.4 *MaxEnt and prawns*

Whilst the analogy between X-ray tomography and fisheries tomography is valid, several differences need to be taken into account when applying MaxEnt to fisheries data.

Tows and associated catches differ from X-rays and absorption density in that observations change the state of the image under investigation; that is, the very thing being measured is removed. This change may manifest itself as a declining measure of CPUE within an area over time, leading to an underestimate of the resource. Factors that will affect the rate of decline would include the intensity of trawling, the time period of data collection, and the biological characteristics of the fish (or prawns) being trawled. Time periods need to be chosen so that enough catch and effort data can be collected to make predictions of stock density in a relatively small area without it being influenced by stock movement.

To determine the optimal data requirements to fit a reliable model, maps of catch were produced using various grid sizes and time intervals for both the shallow and deep EKP sectors. Simple correlation analysis was used to compare catches in grids from one time period to the next. A total of 60 one-day, 20 three-day, and 8 seven-day maps with one, two and three nm grid sizes were generated for each sector within a two-month period.

Fisheries data often contain outliers due to several factors such as unit conversion errors, misidentification of species and data entry errors. Whilst the most obvious outliers are omitted during initial data extraction, there may be outliers on a fine scale – such as a relatively large catch recorded in an area and time where catches are normally very low. To test the robustness of MaxEnt effect on these types of outliers, a correlation analysis was performed. In this analysis, the MaxEnt reconstruction was compared with all the catch data with a number of reconstructions using Jackknife samples. One hundred Jackknife replicates were run with 5, 15 and 30 per cent of tows removed. Various ICF widths were incorporated into each MaxEnt reconstruction. The reconstruction that maximised the evidence term was selected as the most appropriate model.

9.3 Results

Correlation analyses of MaxEnt reconstructions and catch maps were performed by calculating the correlation between original datasets and the same datasets with random tows removed. Grids common to either dataset were used in the analysis. Another option for testing the effect of time period – of using only grids common to both datasets – was rejected as it would lead to a loss in information, especially on a stock moving in space over time.

Figure 9.2 shows the effect on the resulting MaxEnt reconstructions of datasets of one week of EKP data with random tows removed. Mean correlations decrease as more data are removed but still remain relatively high (0.8) even with 30% of the tows removed. Similar results were obtained when a longer time series of two weeks was used. These reconstructions apparently illustrate that MaxEnt is robust to local outliers mentioned previously and also to reduced information. The main effect of removing data is on the accuracy of individual point estimates (Figure 9.3) – standard deviations of point estimates increase with the removal of more data points.

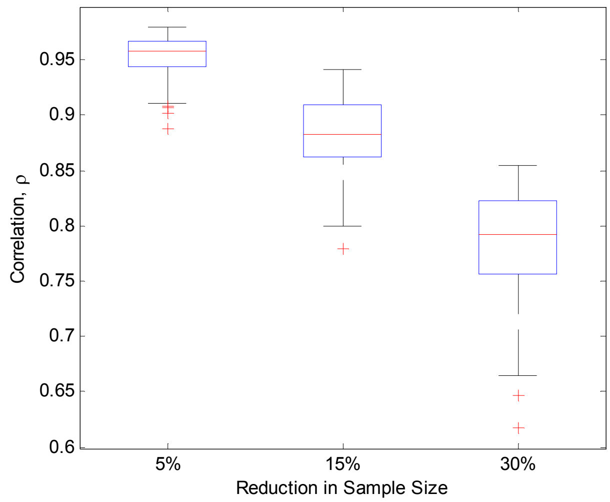


Figure 9.2. Box and whisker plots of correlations of original MaxEnt reconstruction and reconstructions with 5, 15 and 30% of random tows removed, both based on 100 Jackknife sample reconstructions. Data are for a one-week period in December 2003 in the shallow-water EKP fishery and consists of 703 tows.

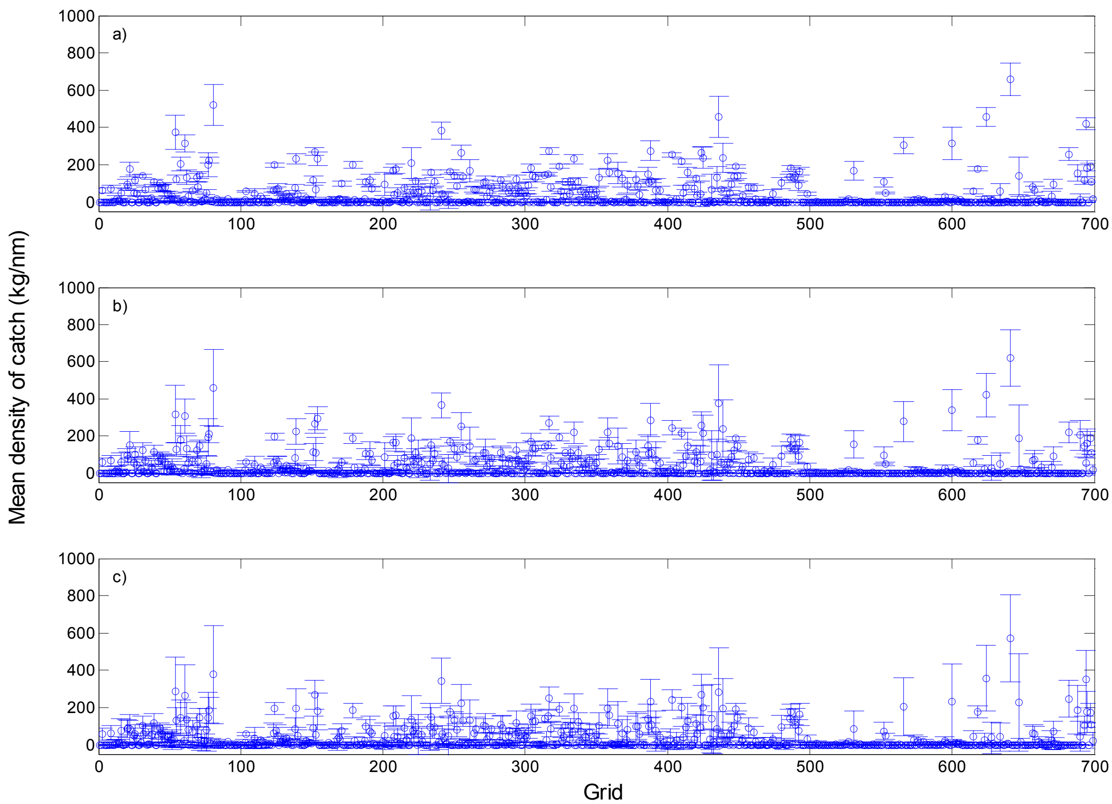


Figure 9.3 Effect on precision and accuracy of a MaxEnt reconstruction by removing random tows. Bars represent one standard error of the mean density of catch (kg/nm) in each grid (*x*-axis): (a) reconstruction with 5% of tows removed, (b) 15% of tows removed, and (c) 30% of tows removed.

There are significant differences in fishing patterns and stock movement between the shallow and deepwater fisheries (Figure 9.4). Correlations between successive time periods in the deepwater fishery are consistently lower than in the shallow-water fishery. Only when lags of up to two or three days occur do both sectors have similar correlation coefficients (Figure 9.4 *b,c*). The highest correlation for the shallow-water fishery is a seven-day dataset using either grid sizes of two or three nm (Figure 9.4 *h,i*). The correlation pattern lasts even after four lag periods. This pattern is the same for three-day comparisons (Figure 9.4 *d,e,f*); however for one-day comparisons there is a sharp drop in correlation for the first three lags before stabilisation is achieved (Figure 9.4 *a,b,c*). For the deepwater fishery, the best combination seems to be for a one-day comparison with a grid size of either two or three nm (Figure 9.4 *b,c*). The correlation drops sharply from a high of 0.6 after a one-day lag to zero after a seven-day lag (Figure 9.4 *c*).

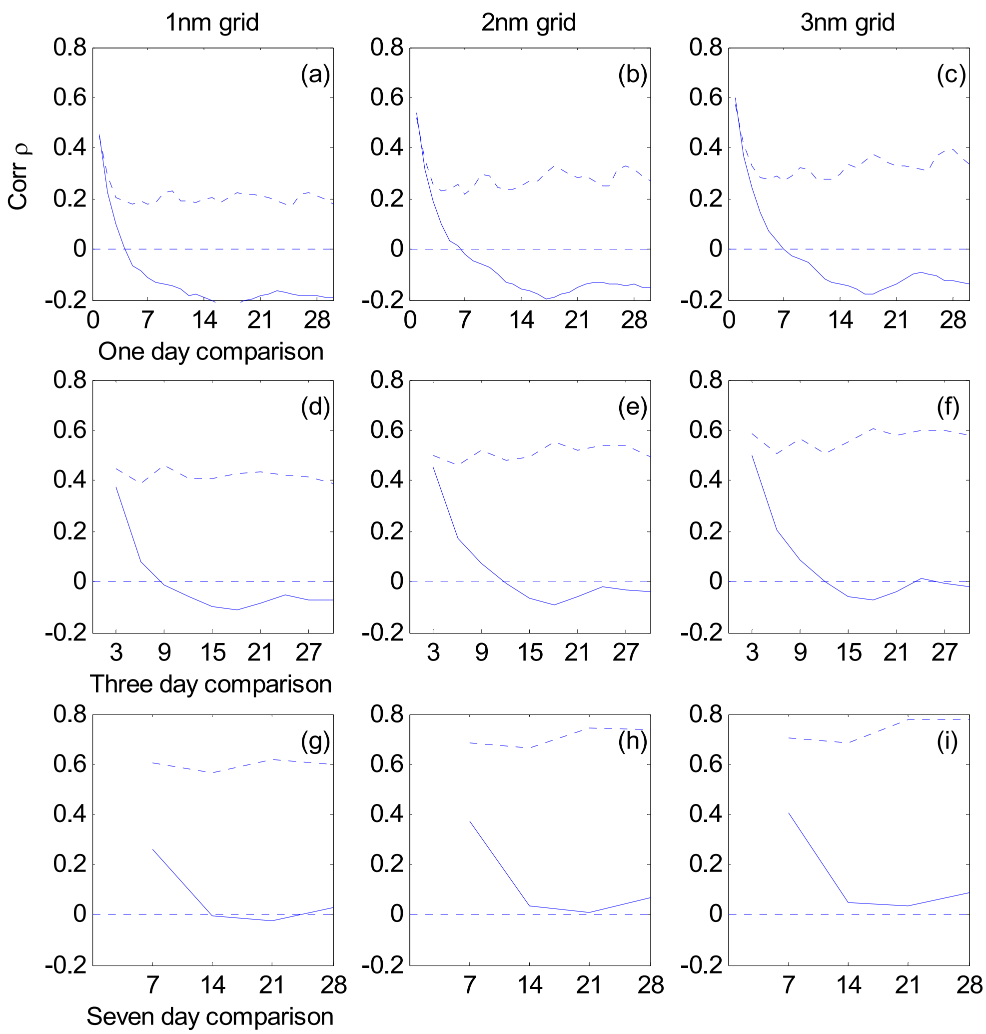


Figure 9.4 Simple Pearson correlations (ρ) of catches in grids from successive time periods (one-, three- and seven-day comparisons) for different grid sizes (one, two and three nm). Solid lines represent comparisons for the deepwater sector and dotted lines the shallow water sector.

Some subjectivity was required to select the optimal ICF width for maps of EKP density. There is marginal gain by incorporating ICF widths greater than six (Table 9.1) – reconstructions produced being overly smooth. Therefore, a cut-off was chosen where there was less than a one per cent improvement in the evidence term. Figure 9.5 shows the effect of ICF widths on the reconstructions in the deepwater sector. A reconstruction with no ICF results in a somewhat speckled image (Figure 9.5 *a*) while increasing the ICF width to five results in definable areas of high and low density (Figure 9.5 *c*). An ICF of width eight tends to over-smooth the image where finer scale detail is not visible (Figure 9.5 *d*).

MaxEnt reconstructions for the deepwater EKP sector using three-day subsets with a grid size of two nm and an ICF width of five are shown in Figure 9.6. Reconstructions using only one-day subsets resulted in a very poor model fit due to a lack of data. Three-day subsets were therefore modelled rather than the one-day subset suggested by the correlation analysis. Correlations for the deepwater sector for three-day subsets were not significantly different from one-day subsets (Figure 9.4 *b, e*). Defined areas of high density are present for most reconstructions (Figure 9.6). The associated map of trawl effort tends to agree with the reconstruction, but the image is blurred (Figure 9.7). It is difficult to infer a definite movement of prawn stock as effort is not consistent over the two-week period. Even so, effort tends to decline in the later periods and is concentrated in the more northerly section of the sector (Figure 9.7 *e, f*).

In the shallow-water sector, there are consistent areas of stock density that are relatively stable over time (Figure 9.8). These areas are much less defined in a map of trawl effort (Figure 9.9), in contrast to the area definition in the deepwater sector. These results suggest that MaxEnt is a reasonably valid method for predicting stock density, especially for a relatively static stock. It is known that in some areas EKP are not historically found such as the patches on the western side of the northern-most island (Figure 9.8 *f*). However, the absence of stock definition may be due to a logbook error by the fisher or an error in the algorithm that defines trawling from the VMS database. Over a fishing season, these errors may become more apparent and can then be rectified.

Table 9.1 Percentage change in evidence for varying ICF widths, using a three-day subset of the deepwater sector.

ICF	Evidence[a]	Per cent
0	-375.9	NA
1	-357.9	4.78
2	-353.2	1.31
3	-347.5	1.61
4	-343.4	1.18
5	-339	1.28
6	-336	0.88
8	-331.5	1.34
10	-330	0.45
12	-328.3	0.51
14	-326.6	0.52
16	-324.7	0.58

[a] Evidence term is calculated as Log_e[Prob(ICFwidth|D, H)].

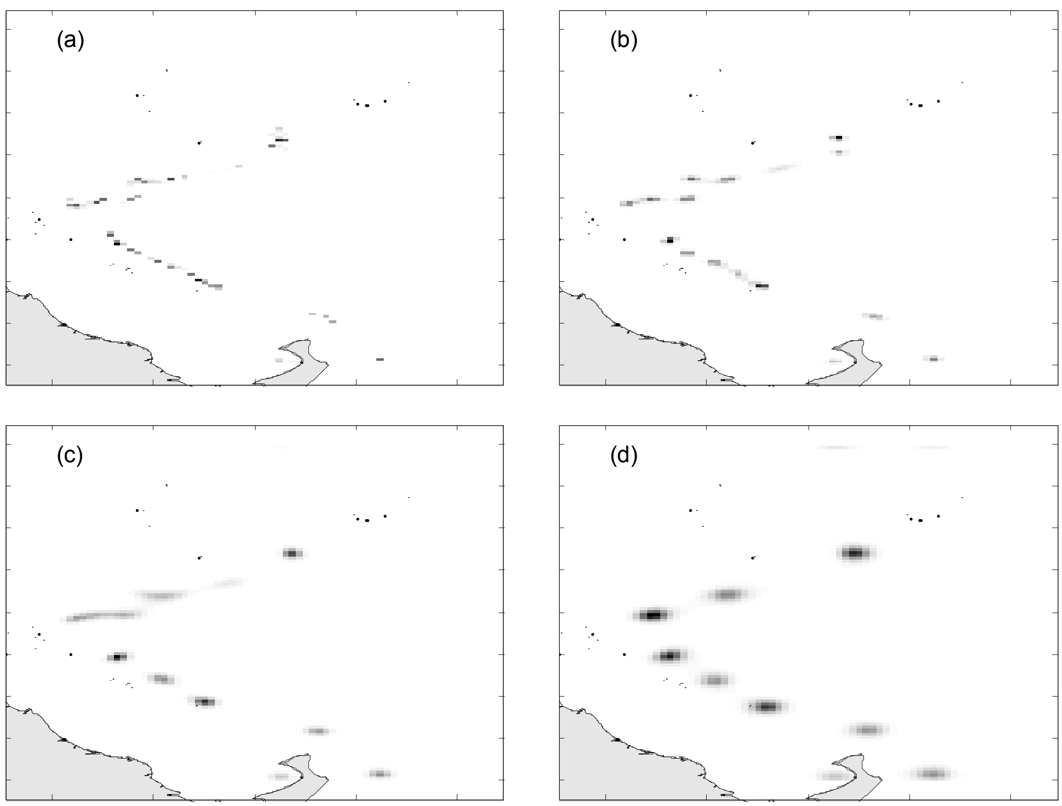


Figure 9.5. MaxEnt reconstructions of EKP distribution in the deepwater sector: (a) No ICF (b) ICF width of 3 (c) ICF width of 5 and (d) an ICF width of 8.

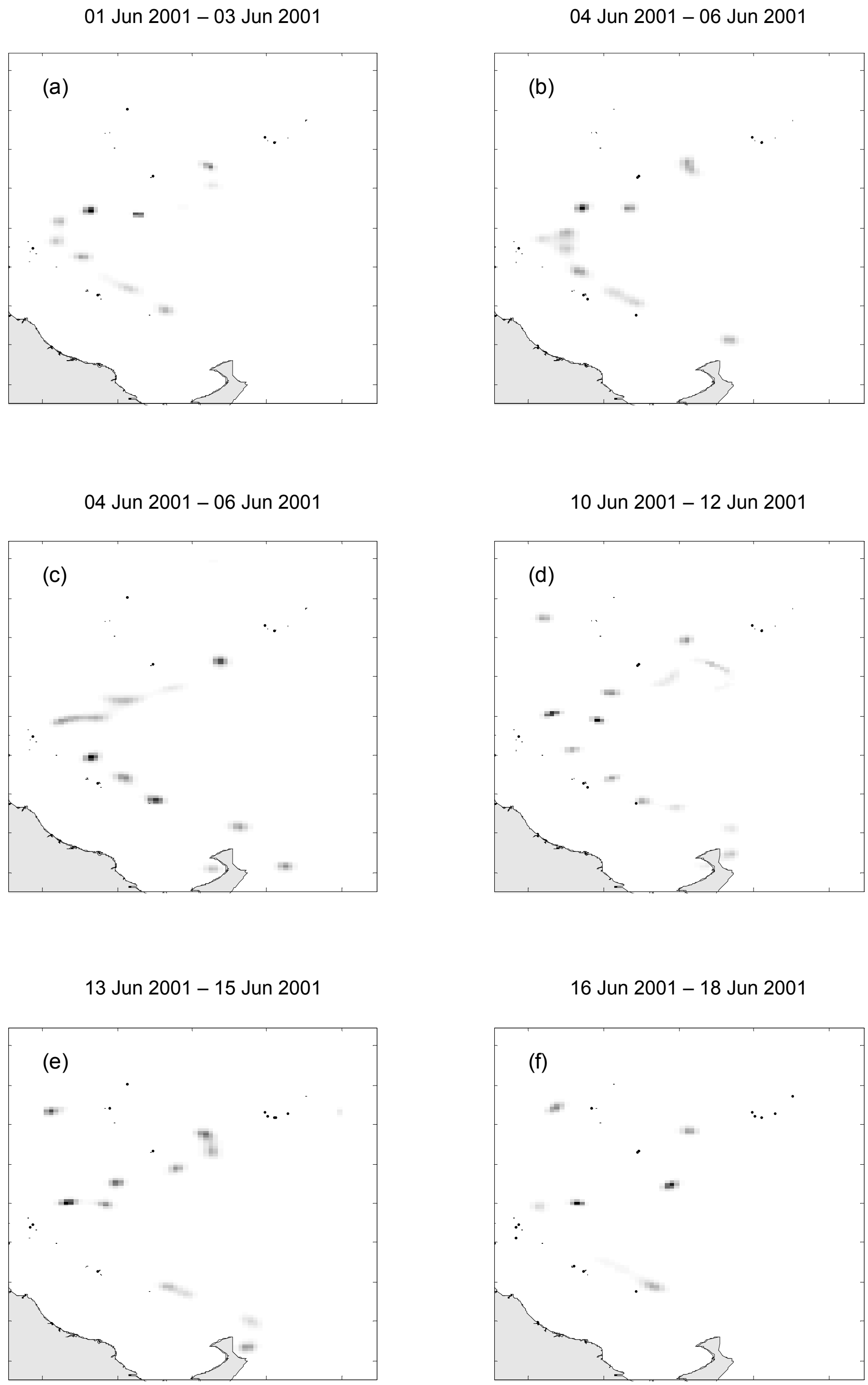


Figure 9.6. MaxEnt reconstruction of deepwater EKP using an ICF width of five for three-day subsets from 1 June 2001 to 16 June 2001. Light-coloured grids represent low-resource density, and dark-coloured grids represent high-resource density.

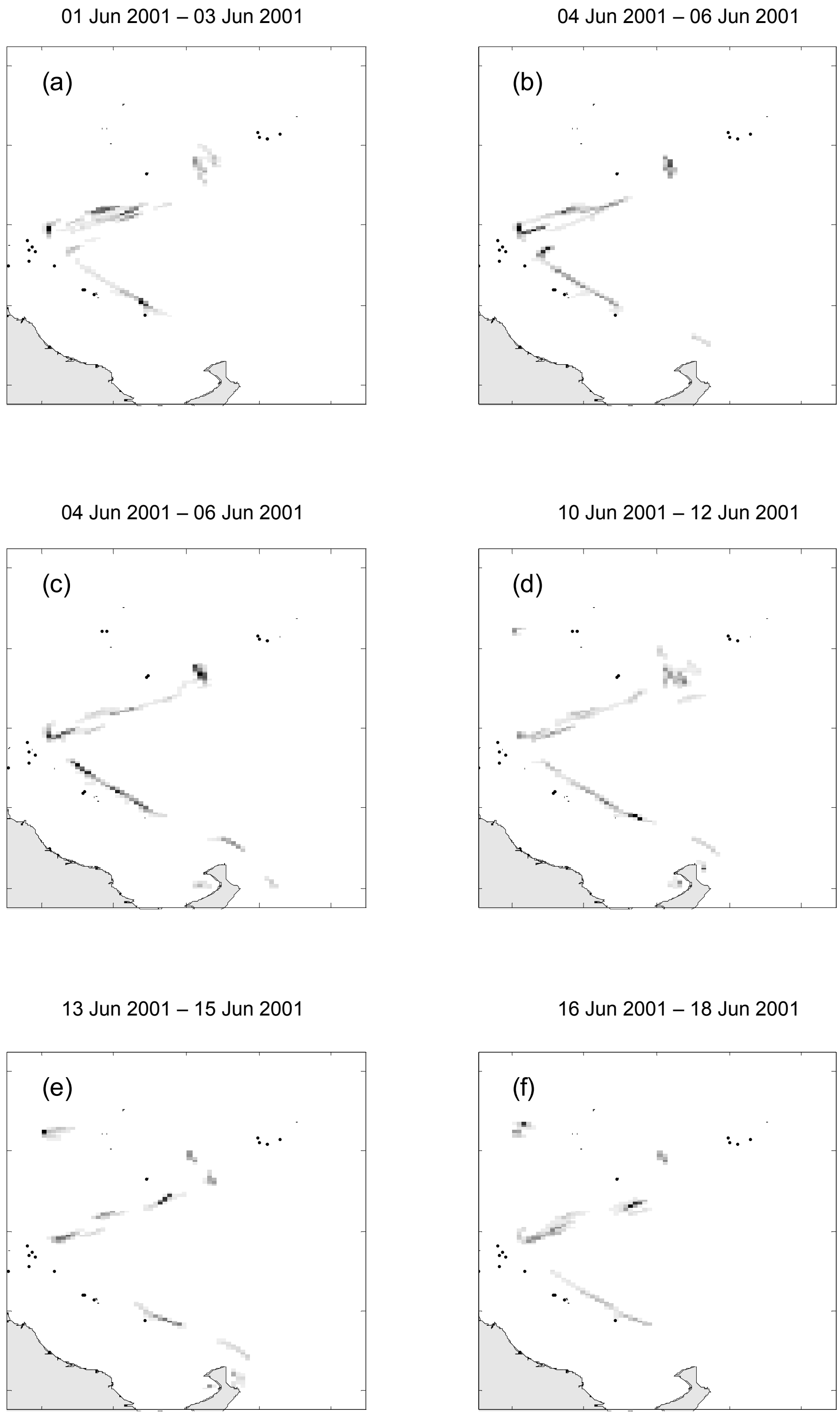


Figure 9.7. Trawl effort (measured as nm tows) in the deep-water EKP for three day subsets from 1 January 2001 to 18 January 2001. Grids with less than five boats operating have been removed for confidentiality reasons.

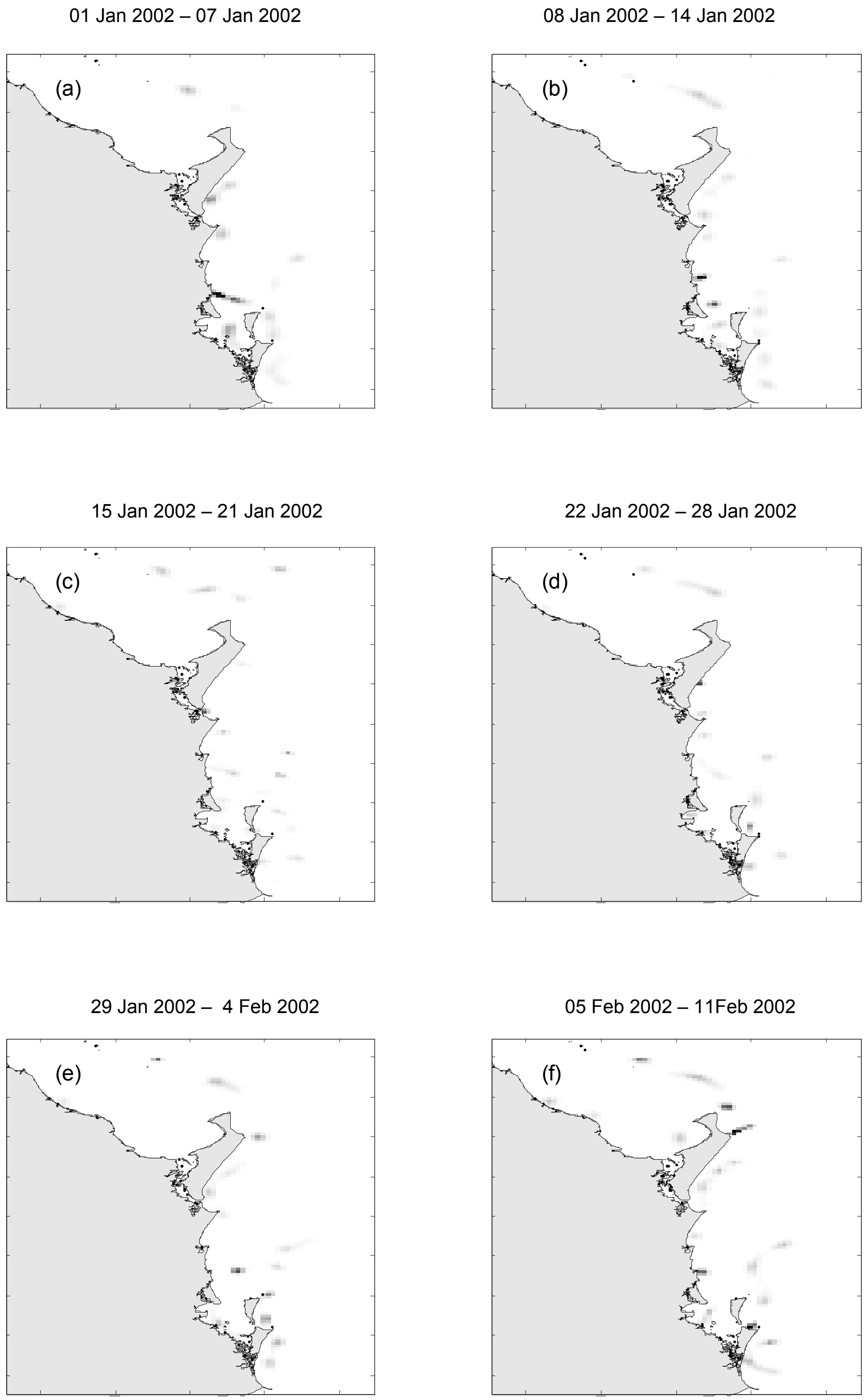


Figure 9.8. MaxEnt reconstruction of shallow-water EKP using an ICF width of five for seven-day subsets from 1 January 2002 to 5 February 2002. Light-coloured grids represent low-resource density, dark-coloured grids high-resource density.

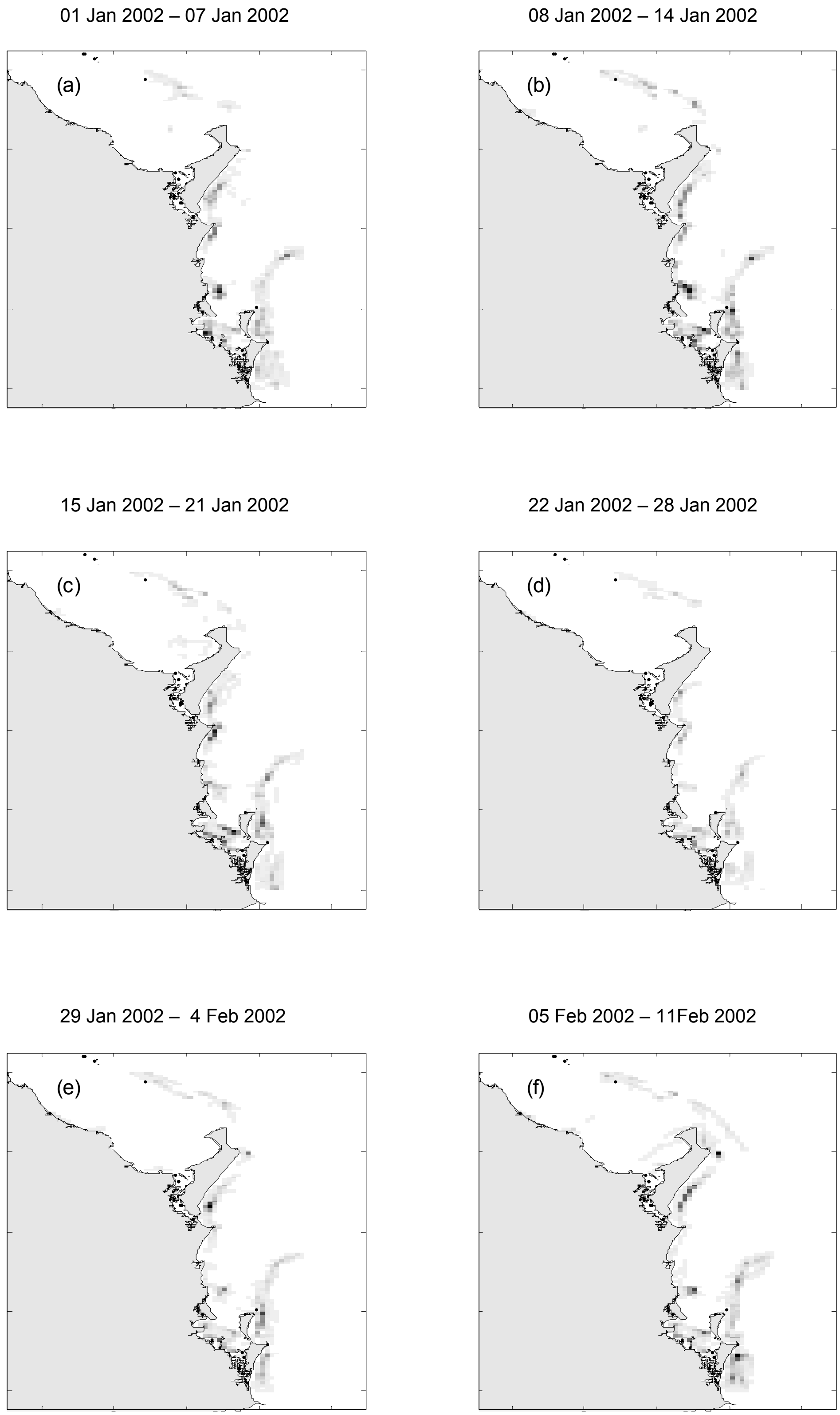


Figure 9.9. Trawl effort (measured as nm tows) in the shallow-water EKP for seven day subsets from 1 January 2002 to 11 February 2002. Grids with less than five boats operating have been removed for confidentiality reasons.

9.4 Discussion

The use of an intrinsic correlation function within the MaxEnt formulation allows detail to be ‘borrowed’ from surrounding grid cells to create a more realistic image of stock distribution. When doing so however, it is important to ensure that widths do not exceed the likely boundaries of the stock which may have been present in the deepwater sector. Brierley *et al*. (2003) incorporated multiple scales so that the reconstruction could either be smooth in adjacent areas with similar density, or sharp in areas with defined aggregations. The use of such scales will be investigated, especially for the deepwater sector where there are defined narrow ‘highways’ following depth contours and in areas where the stock is less aggregated. Whilst the evidence term may suggest incorporating larger ICF widths into the MaxEnt model, it is essentially up to the operator to make a subjective decision as to what is the most feasible ICF width. The main benefit found in this study for using an ICF was to allow the use of smaller grid sizes to map stock distribution. Vignaux *et al*. (1998) used grid cells of 8 × 8 nm to gain reasonable degree of detail in their MaxEnt reconstructions of the hoki fishery in New Zealand. Incorporating an ICF into MaxEnt may well allow the authors to reduce their grid size, as hypothesised by Vignaux *et al*. (1998).

Figure 9.10 shows the difference between a MaxEnt reconstruction using an ICF width of five and a map of standardised CPUE. The MaxEnt reconstruction clearly shows areas of high catch density, whereas density estimates from the map of standardised CPUE are much lower and more widely dispersed. Some estimates of MaxEnt-derived catch density seem overly high and warrant further attention. The incorporation of a non-flat prior consisting of average catch per grid from a previous time period might aid in forcing the model to predict more realistic densities (S. Gull, pers. comm.). Altering the number of scales may also lead to improvements. Another issue to be addressed in later work is ensuring that the model is set to zero on land. In our study, there were some instances where values for grids on dry land were given, and where catch was only a few kilometres from the coast for the shallow-water sector.

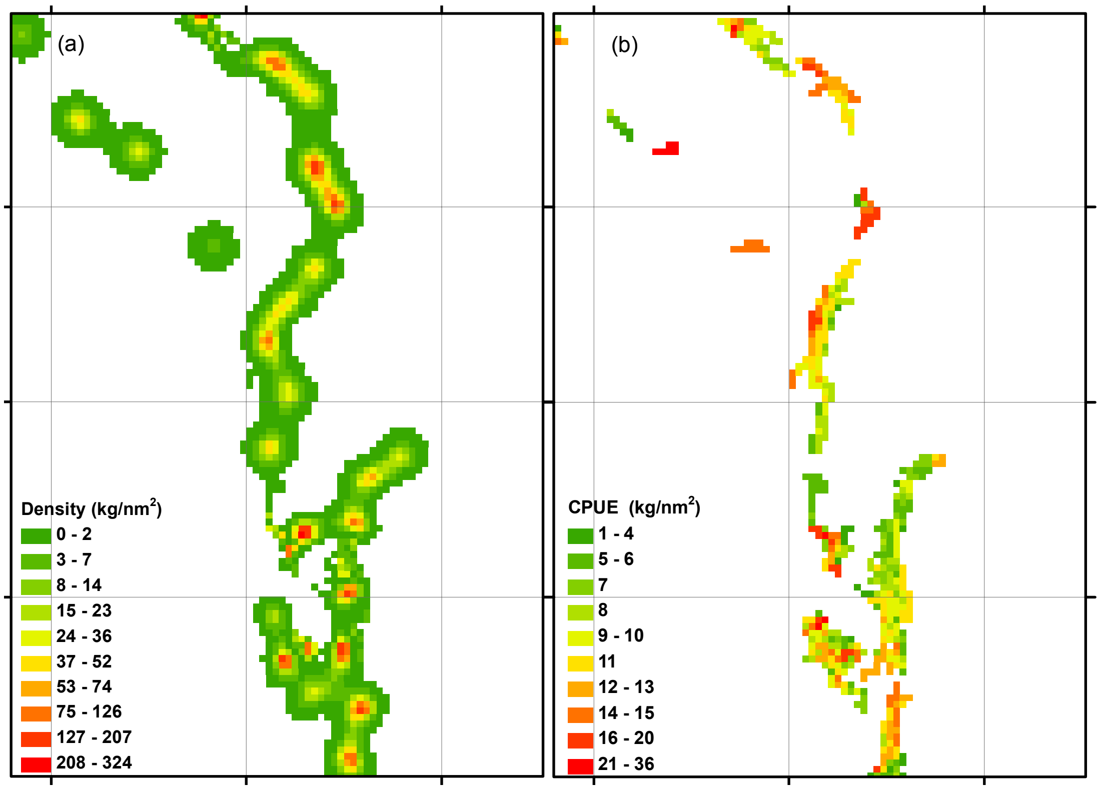


Figure 9.10. Comparisons between a MaxEnt reconstruction using an ICF width of five (a), and a map of CPUE (b) on data gathered from a week in January 2002, in the shallow-water EKP. Identifying features such as land and coordinates have been omitted for confidentiality.

The use of VMS-derived effort data has provided fisheries resource managers with a valuable tool for mapping the distribution of fishing effort at very fine scales. However, this chapter has shown that using MaxEnt to predict resource density is a better tool for mapping catch density at very fine scales than are maps of ‘blurred’ CPUE. Whilst the accuracy of predicted densities may be less than desired, MaxEnt does at least consistently predict the likely location and spatial extent of the resource. This information may be useful for a number of applications including defining spawning aggregations, aiding in designing monitoring surveys, and predicting stock movement and depletion rates. Further improvements may eventually lead to the estimation of the overall stock biomass.

10 POTENTIAL IMPROVEMENTS TO THE DATA REQUIRED FOR STOCK ASSESSMENT AND THE EVALUATION OF REFERENCE POINTS

Mai Tanimoto and Rick Officer

10.1 Introduction

Current assessment models for estimating stock biomass are calibrated using commercial logbook catch and effort data. These data are used to create standardised CPUE indices of abundance, which aim to remove variations in the CPUE data not attributed to the abundance of the population (e.g. area fished, season, and fishing power of vessels). Unfortunately, the current standardisation models developed for Queensland's east coast fisheries may generate hyperstable indices of abundance as they are all based on a relatively coarse spatial scale (30 nm × 30 nm CFISH grids). At this scale the local depletion of stocks may be obscured. Improvements to the standardisation of these indices are therefore critical to improving the precision of stock assessment outputs and to providing greater certainty in management advice.

The *Fisheries (East Coast Trawl) Management Plan 1999* specifies catch rate limit reference points for the principal targeted species of Queensland's east coast trawl fisheries, including EKPs, bay prawns, Moreton Bay bugs, redspot king prawns, saucer scallops and tiger prawns. The review events were defined to be triggered when catch rate (CPUE) for the targeted species fell below 70% of those from 1988 to 1997 within the recruitment or spawning periods. The review periods specified for EKPs are November to February inclusive (recruitment period) and May to August inclusive (spawning period).

By using the EKP fishery as a case study, this chapter incorporated finer scale spatial information of catch and effort into the existing CPUE standardisation model. We investigated how this VMS-derived information improved the standardised relative index of abundance and whether or not these new standardised CPUE triggered the review events for this fishery.

10.2 Method

10.2.1 *Data*

The daily logbook catch and effort (CFISH) records and VMS data were linked by the TrackMapper software. The trawl signature of the EKP fishery (developed in Chapter 7) was applied to retrieve the monthly catch and effort records of the fishery for the period between January 2001 to December 2004. In order to obtain the precise spatial distribution of fishing effort, all records were obtained at the finest spatial resolution available of 1 nm × 1 nm cell.

10.2.2 *Statistical analysis*

Estimation of effort exerted and its spatial distribution, 1988–2000

The spatial distribution of monthly fishing effort was incorporated into the catch rates standardisation model of the EKP fishery from 1988 to 2004 in terms of a) $\mathrm{Density}$ (total fishing hours) and b) Area (count of the number of one-minute cells fished within each 30-minute grid). As we have only VMS-derived catch and effort data since 2001, individual monthly estimates of fished area in each CFISH 30-minute grid for the entire time series from 1988–2004 is not possible. Fishing hours and area fished were therefore estimated for each month during the missing period (January 1988 – December 2000) by using a generalised linear model (GLM). The monthly fishing effort (number of days fished) obtained from the CFISH database was likely to indicate the trend of fishing hours and areas, and was therefore included in the model. After several exploratory analyses, the best prediction models were determined as follows:

$$\left.\begin{array}{l}\log_{\mathrm{link}}(\mathrm{Density}) \\ \log_{\mathrm{link}}(\mathrm{Count})\end{array}\right\} = \mathrm{month} + \mathrm{grid} + \mathrm{month.logmdays} + \mathrm{grid.logmdays} + \varepsilon$$

Where $\mathrm{Density}$ is the monthly total hours fished within 30-minute grids, Count is the monthly total number of one-minute cells fished within 30-minute grids, grid is each CFISH 30-minute grid, $\mathrm{logmdays}$ is the log-transformed number of days fished in each month, and ε is the Poisson error term. In order to directly evaluate the impact of including VMS-derived factors in the standardisation, this study was restricted to the records from the 14 selected grids that have previously been used in analyses of fishing power in the Queensland EKP fishery (O'Neill and Leigh 2006). O'Neill's analyses selected only the 30-minute grids where a clear distinction in the depth of trawling existed. Grids characterised by a mixture of deep and shallow trawling for EKP were excluded.

Standardisation of CPUE, 1988–2004

Once the predicted monthly fishing hours and area fished between 1988 and 2000 were obtained, these data together with raw VMS-derived effort data between 2001 and 2004 were combined with daily CFISH logbook data (1988–2004) to conduct the standardisation of CPUE. This study used the standardisation model for the Queensland EKP fishery developed by O'Neill and Leigh (2006): a linear mixed model using the method of residual maximum likelihood (REML) assuming normally distributed errors. The statistical software package GenStat Version 8 (GENSTAT 2005) was used to undertake the analysis.

The response variable was based on the daily catches of EKPs by individual vessels within each 30-minute CFISH grid. In order to examine whether VMS-derived information improved the standardisation of CPUE, alternative models were run with and without VMS information.

The models were defined as follows:

Model A (with VMS) $\log_e\left(C_{vayml}\right) = \beta_0 + \sum \boldsymbol{\beta}_1\mathbf{X}_1 + \sum \boldsymbol{\beta}_2\mathbf{X}_2 + \sum \boldsymbol{\beta}_3\mathbf{X}_3 + \beta_4 X_4 + \sum \boldsymbol{\beta}_{VMS}\mathbf{X}_{VMS} + \boldsymbol{\gamma}\mathbf{Z} + \varepsilon$

Model B (without VMS) $\log_e\left(C_{vayml}\right) = \beta_0 + \sum \boldsymbol{\beta}_1\mathbf{X}_1 + \sum \boldsymbol{\beta}_2\mathbf{X}_2 + \sum \boldsymbol{\beta}_3\mathbf{X}_3 + \beta_4 X_4 + \boldsymbol{\gamma}\mathbf{Z} + \varepsilon$

where C_{vayml} was the daily catch of the v^{th} vessel in grid a, during fishing year y, month m and lunar cycle l; β_0 was a scalar intercept parameter to be estimated; $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, β_4 and $\boldsymbol{\beta}_{VMS}$ were vector (except β_4 being as a scalar**)** parameters to be estimated for abundance, catchability (fishing power), lunar phase, saucer scallop catches and VMS information respectively; $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, X_4 and $\mathbf{X}_{VMS}$ were the corresponding data; the γ was a random term for vessel $\mathbf{Z}$; and ε was the normal error term. $\mathbf{Z}$ indicates which daily catches belonged to each vessel (record-number). The biomass vector $\boldsymbol{\beta}_1$ was expressed by the two-way interaction effects of different abundance terms including fishing grids, fishing years and months. Note that the fishing year for the EKP fishery starts in November and finishes in October (O'Neill and Leigh 2006). The catchability of prawns $\boldsymbol{\beta}_2$ was represented by the vector of capture system variables including different vessel characteristics, navigation equipment, bycatch reduction devices and trawl net configurations. This component of the model was the exclusive focus of interpretation to calculate annual changes in fishing power. The parameters $\boldsymbol{\beta}_3$ represented lunar cycles for luminance and luminance advanced seven days. The VMS parameters $\boldsymbol{\beta}_{VMS}$ were represented by the two two-way interactions of hours fished (log-transformed Density) and fishing grid (logdensity.grid), and area fished (log-transformed Count) and fishing grid (logarea.grid). The result from the abundance vector ($\boldsymbol{\beta}_1$), specifically the interaction between year and month terms, was used to calculate standardised monthly catch rates. Details of the original standardisation model (Model B without VMS information) are available in O'Neill and Leigh (2006: pp 21-25).

The results from Models A and B were compared in terms of a) statistical significance of the VMS parameters (α = 0.05), and b) the overall model goodness of fit. Model goodness of fit was examined by Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). It is again important to note that raw VMS information were not available before 2001, and true influence of the VMS parameters on the standardisation model were unlikely to be detected if we used the 17 years of CFISH logbook data (1988–2004) with 13 years of predicted VMS effort from 1988 to 2000. To test the influence of the VMS parameters, standardisation models A and B were applied to the data for two time periods: a) from 2001 to 2004, and b) from 1988 to 2004. The first period was used to investigate the significance of the VMS parameters on the model, and the second period was used to obtain the time series of standardised CPUE between 1988 and 2004.

10.3 Results

10.3.1 *VMS-derived fishing hours and area*

The monthly VMS-derived fishing hours and areas fished in 14 selected CFISH grids are shown in black lines in Figure 10.1 and Figure 10.2 respectively. In general, trends in fishing hours vary within and across the CFISH grids, indicating that there are different temporal and spatial distributions of fishing effort in the EKP fishery. Seasonal patterns of fishing hours were found in most of the CFISH grids, but their peak seasons were variable among grids. For example, peaks of fishing effort (hours fished) in grids W34 and W36 were found in early summer (Nov–Dec), whilst grids X35 and X36 had peaks during late autumn (May–June).

The trends in fishing area (count of one-minute cell) also varied across the CFISH grids. Temporal distributions were relatively consistent in grids U28, U29, V30, X35 and X36 (about 200 cells per month), while strong seasonal variation were found in grids V31, W28, W34 and W36. Their seasonal peaks also vary among these grids.

Both of the GLM models were well fitted to the raw VMS effort and area data as illustrated by red lines in Figure 10.1 and Figure 10.2.

Figure 10.3 and Figure 10.4 show the time series of predicted (1988–2000) and raw (2001–2004) VMS effort and area that were incorporated into the standardisation model. The long-term trends in predicted effort and area were relatively stable throughout the years.

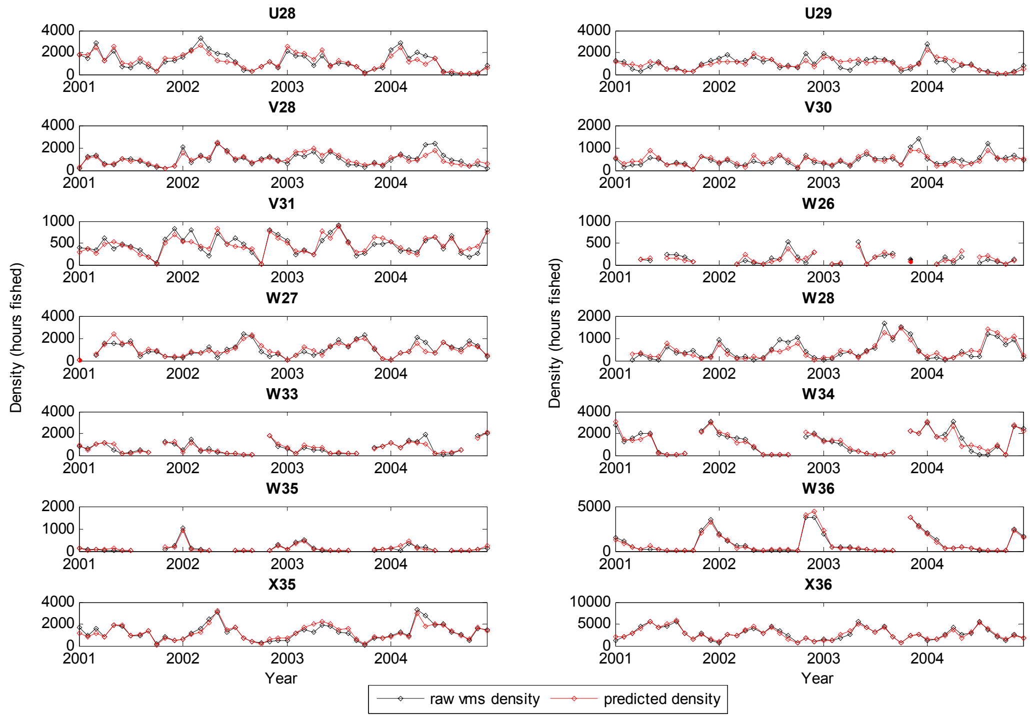


Figure 10.1 Trends of raw and predicted VMS-derived fishing effort (hours fished) in 14 selected CFISH grids (2001–2004). The predicted values from the GLM model were well fitted to the raw VMS data.

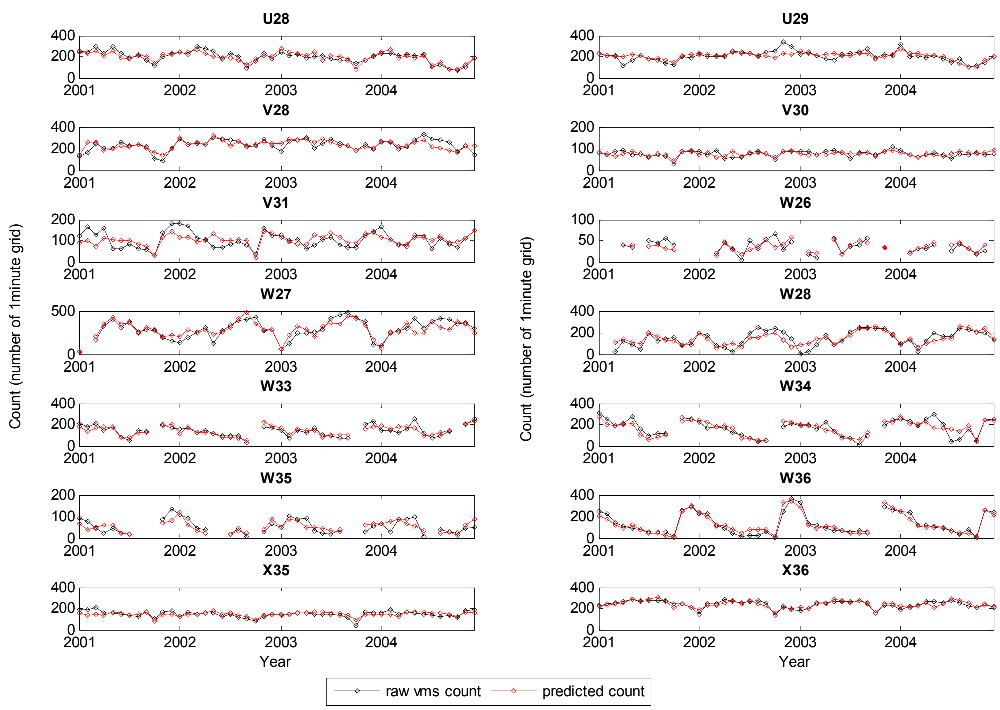


Figure 10.2. Trends of raw and predicted VMS-derived fishing area (count of one-minute cells) in 14 selected CFISH grids (2001–2004). The predicted values from the GLM model were well fitted to the raw VMS data.

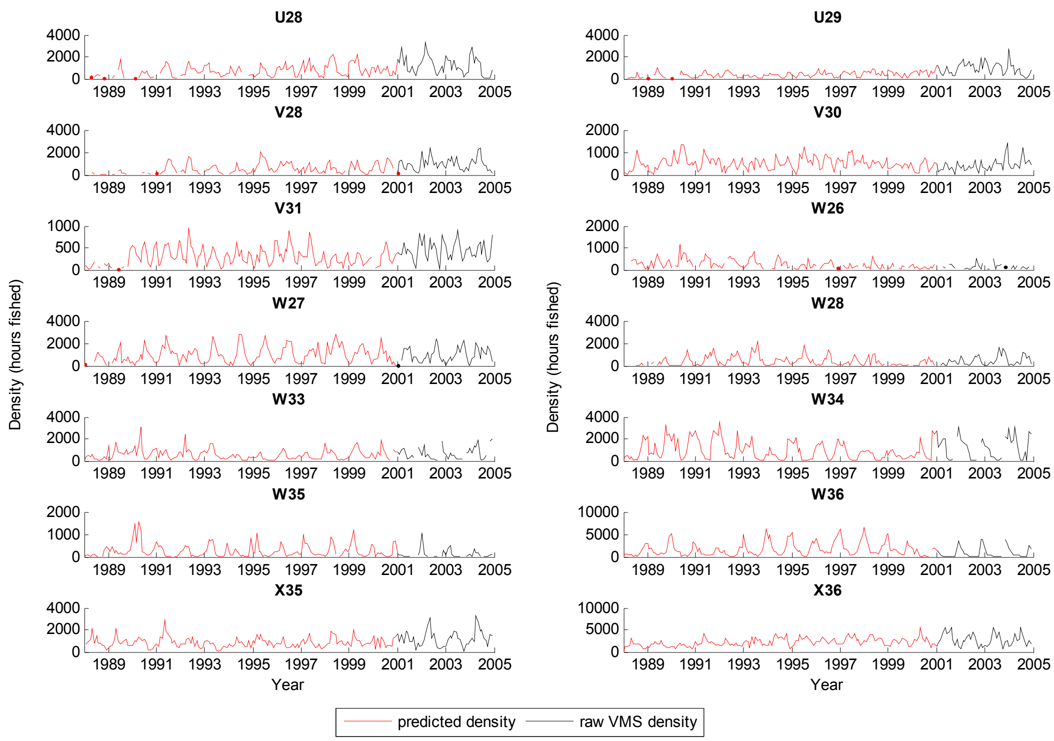


Figure 10.3. The time series of VMS-derived monthly fishing hours from 1988–2004, which were incorporated into the standardisation model.

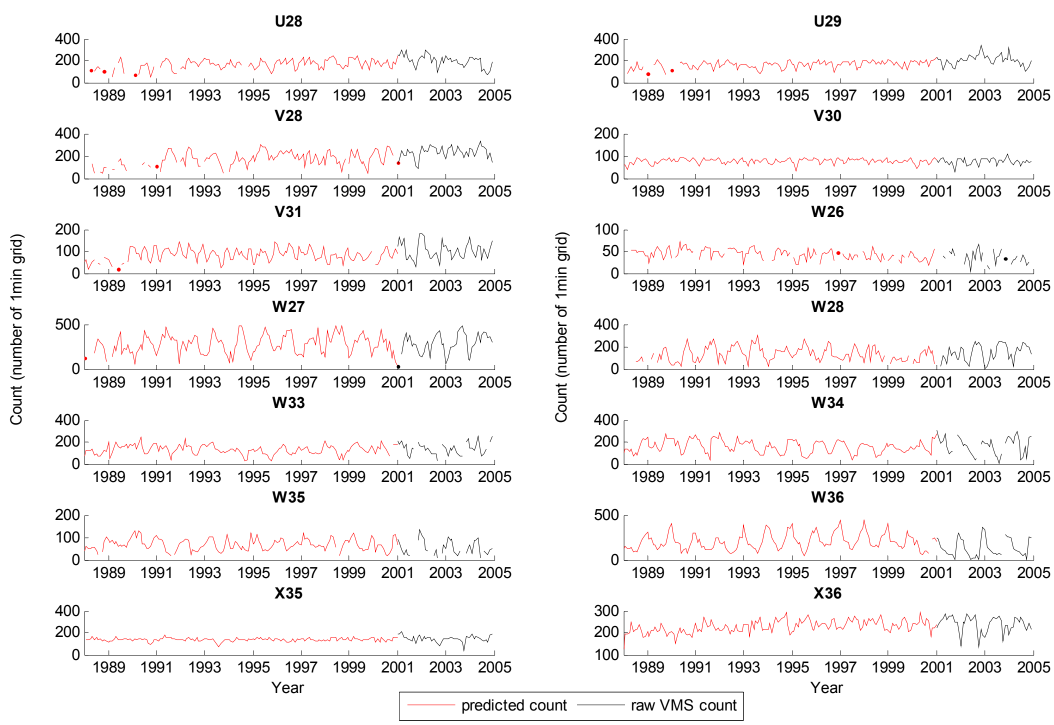


Figure 10.4. The time series of VMS-derived monthly area fished (count) from 1988–2004, which were incorporated into the standardisation model.

10.3.2 *Influence of VMS-derived parameters*

Results of the linear mixed model (REML) applied to CFISH data with (Model A) and without (Model B) VMS information between 2001 and 2004 are shown in Table 10.1. The Wald statistics of VMS parameters, logarea.grid and logdensity.grid, were both significantly different from zero ($p < 0.001$), indicating that the effect of fishing area and hours varies between the CFISH grids (Table 10.1). Results of the standardisation model applied to the entire CFISH data series (1988 to 2004) also indicated that VMS parameters were significant (Table 10.2).

Table 10.1. Results of the linear mixed models (REML) using data from 2001–2004 (the period when raw VMS data were available) calculated with and without the inclusion of VMS-derived information. Wald statistics were calculated by dropping each fixed term from the full exploratory model.

	Random term	
	Model A (with VMS)	**Model B (without VMS)**
Estimated Variance Components (S.E.) Record_number	0.163 (0.0255)	0.1556 (0.0245)

	Residual term	
	Model A (with VMS)	**Model B (without VMS)**
Residual variance (S.E.)	0.234 (0.0021)	0.237 (0.0021)
Deviance (-2Log-likelihood)	-7399.54	-7291.74
Residual degrees of freedom (d.f)	24475	24503

		Wald Statistics		**χ^2 probability**	
Fixed terms	**d.f.**	**Model A (with VMS)**	**Model B (without VMS)**	**Model A (with VMS)**	**Model B (without VMS)**
logarea.grid (VMS parameter)	14	106.17	~	<0.001	~
logdensity.grid (VMS parameter)	14	156.13	~	<0.001	~
boards	3	23.7	22.97	<0.001	<0.001
brdted	1	0.01	0.09	0.941	0.764
compmap	1	2.89	4.2	0.089	0.04
fishyear.grid	39	356.6	387.95	<0.001	<0.001
fishyear.month	25	336.46	341.24	<0.001	<0.001
ggear4	4	8.28	8.67	0.082	0.07
logchain	1	18.45	16.85	<0.001	<0.001
logmesh	1	9.76	9.14	0.002	0.003
logmetres	1	34.61	37.07	<0.001	<0.001
logscallops	1	535.38	582.92	<0.001	<0.001
lunar	1	51.93	51.22	<0.001	<0.001
lunar_adv	1	65.85	72.48	<0.001	<0.001
month.grid	134	1294.3	2206.43	<0.001	<0.001
nettype	3	17.01	16.91	<0.001	<0.001
sonar	1	1.89	1.12	0.169	0.29
tryyesno	1	0.04	0.03	0.845	0.87

Table 10.2. Results of the linear mixed models (REML) using data from 1988–2004 calculated with and without the inclusion of VMS-derived information. Wald statistics were calculated by dropping each fixed term from the full exploratory model.

	Random term	
	Model A (with VMS)	**Model B (without VMS)**
Estimated Variance Components (S.E.) Record_number	0.1896 (0.021)	0.1951 (0.022)

	Residual term	
	Model A (with VMS)	**Model B (without VMS)**
Residual variance (S.E.)	0.302 (0.0015)	0.306 (0.0015)
Deviance (-2Log-likelihood)	-9458.9	-8743.67
Residual degrees of freedom (d.f)	82279	82307

		Wald Statistics		**χ^2 probability**	
Fixed terms	**d.f.**	**Model A (with VMS)**	**Model B (without VMS)**	**Model A (with VMS)**	**Model B (without VMS)**
logarea.grid (VMS parameter)	14	83.48	~	<0.001	~
logdensity.grid (VMS parameter)	14	384.2	~	<0.001	~
boards	3	27.54	33.53	<0.001	<0.001
brdted	1	152.76	172.72	<0.001	<0.001
compmap	1	7.39	7.83	0.007	0.005
fishyear.grid	208	2490.17	2767.89	<0.001	<0.001
fishyear.month	168	2647.01	2741.95	<0.001	<0.001
ggear4	4	107.52	116.22	<0.001	<0.001
logchain	1	31.38	30.83	<0.001	<0.001
logmesh	1	212.64	222.95	<0.001	<0.001
logmetres	1	273.83	275.75	<0.001	<0.001
logscallops	1	1281.85	1276.79	<0.001	<0.001
lunar	1	46.12	45.92	<0.001	<0.001
lunar_adv	1	214.52	218.51	<0.001	<0.001
month.grid	143	2575.41	5466.26	<0.001	<0.001
nettype	4	156.9	169.33	<0.001	<0.001
sonar	1	6.15	6.5	0.013	0.011
tryyesno	1	25.12	26.71	<0.001	<0.001

Table 10.3 shows the VMS-derived parameter estimates calculated from Model A (with VMS information) using two data periods. Whilst the effects of the VMS-derived parameters were generally similar for both periods of data, parameter estimates from 2001–2004 data were considered to be more reliable as raw VMS-derived records were available during this period. In general, VMS-derived fishing hours (logdensity.grid) was found to have a positive effect on catch rates, in contrast to VMS-derived fishing area (logarea.grid), which tends to have a negative effect. These results have the following implications:

- Increased fishing hours increases EKP catches in most of the selected CFISH grids (except in U29, V31, W26)
- Increased fishing area decreases EKP catches in grids U29, W27, W33, W34, W36, X35 and X36 (see columns '01–04' in Table 10.3).

Table 10.3. Estimates of the VMS-derived parameters in Model A (with VMS) applied for two CFISH data periods (2001–2004 and 1988–2004). The values in brackets are the standard error of estimated values.

	logdensity.grid		logarea.grid	
Grid	01-04	88-04	01-04	88-04
U28	0.15 (0.05)**	0.19 (0.05)**	-0.08 (0.16)	-0.19 (0.15)
U29	0.00 (0.04)	0.05 (0.03)	-0.3 (0.09)**	-0.22 (0.1)*
V28	0.19 (0.04)**	0.09 (0.04)*	-0.19 (0.11)	-0.26 (0.1)*
V30	0.33 (0.08)**	0.19 (0.04)**	0.19 (0.21)	-0.03 (0.19)
V31	-0.43 (0.15)**	-0.12 (0.07)	0.14 (0.22)	0.33 (0.13)*
W26	0.19 (0.13)	0.07 (0.09)	-0.07 (0.42)	-0.12 (0.26)
W27	0.13 (0.05)**	0.14 (0.04)**	-0.21 (0.09)*	-0.21 (0.08)*
W28	0.12 (0.06)*	0.16 (0.04)**	-0.14 (0.08)	-0.16 (0.07)*
W33	0.25 (0.04)**	0.26 (0.03)**	-0.42 (0.1)**	-0.11 (0.08)
W34	0.10 (0.04)*	0.33 (0.03)**	-0.26 (0.09)**	-0.31 (0.08)**
W35	0.27 (0.07)**	0.21 (0.04)**	0.07 (0.14)	-0.08 (0.12)
W36	0.12 (0.04)**	0.21 (0.03)**	-0.45 (0.1)**	-0.24 (0.08)**
X35	0.19 (0.06)**	0.19 (0.03)**	-0.56 (0.18)**	-0.43 (0.12)**
X36	0.13 (0.06)*	0.15 (0.04)**	-0.81 (0.16)**	-0.55 (0.14)**

** = $p \le 0.01$; * = $0.01 < p \le 0.05$

10.3.3 *Goodness of fit*

Table 10.4 shows the model goodness of fit measures (AIC and BIC) calculated for Models A and B for the extended and recent data periods. The goodness of fit measures indicated that the inclusion of VMS parameters improved the standardisation model for the extended data period (1988–2004). However, goodness of fit measures calculated for the standardisation model of the recent data period (2001–2004) are considered to be more reliable as raw (rather than calculated) VMS records were available during this period. The AIC indicated that the inclusion of VMS parameters improved the standardisation model for the recent data period, but the AIC is known to be biased, favouring more complex models when the sample size is large (> 1000) (SHONO 2005). Whilst the BIC indicated that the exclusion of VMS parameters resulted in a marginally better standardisation model, there was little practical difference between the two models. These results indicate that the addition of the two VMS parameters (logarea.grid and logdensity.grid) resulted in marginal improvement in the standardisation model.

Table 10.4. Results of model goodness of fit measures AIC and BIC.

Data period	Model	Residual SS	Number of parameters	Number of observations	Residual d.f.	AIC	BIC
2001-2004	Model A (with VMS)	5713.8	247	24751	24475	**24969.0**	26973.8
	Model B (without VMS)	5774.0	219	24751	24503	25171.0	**26948.6**
1988 -2004	Model A (with VMS)	24829.5	569	82890	82279	**83417.0**	**88723.1**
	Model B (without VMS)	25096.7	541	82890	82307	84246.4	89291.4

10.3.4 *Standardised catch rate*

Standardised monthly catch rates predicted from Models A and B using data for the entire CFISH period (1988–2004) are shown in Figure 10.5 and Figure 10.6. The trends shown in Figure 10.5 are given on a relative scale to allow comparison of the two models. Figure 10.6 compares the absolute standardised catch rates calculated for each model with the 70% catch rate review period within each fishing year. Note that the dotted red and blue lines illustrate the limit reference points for the recruitment period (November to February inclusive) and the spawning period (May to August inclusive) respectively. In general, similar trends were obtained from the two models, but the peaks of Model A are slightly lower than that of Model B (Figure 10.5). In terms of review events, both models produced the same conclusion, that is, standardised monthly catch rates between 1998 and 2004 were all above the corresponding 70% CPUE reference points except in January 2000.

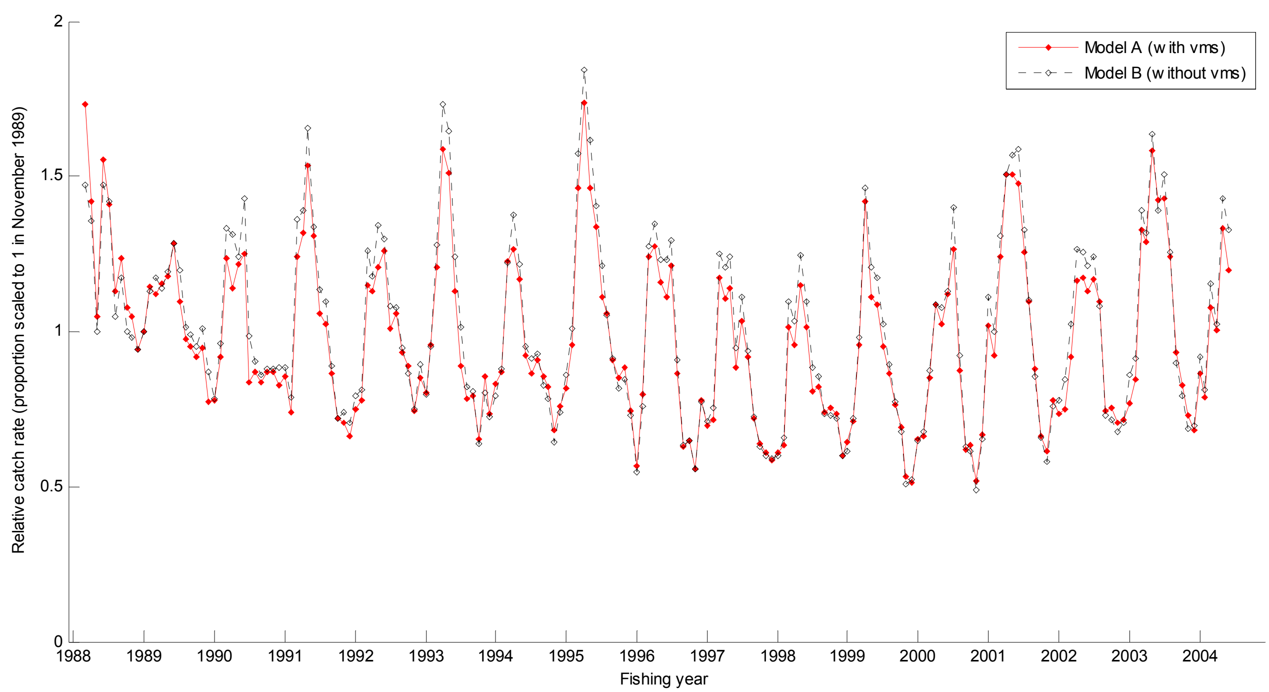


Figure 10.5. Relative standardised catch rates from Model A (with VMS, red line) and Model B (without VMS, black line). Both series are scaled relative to a value of 1 in November 1989.

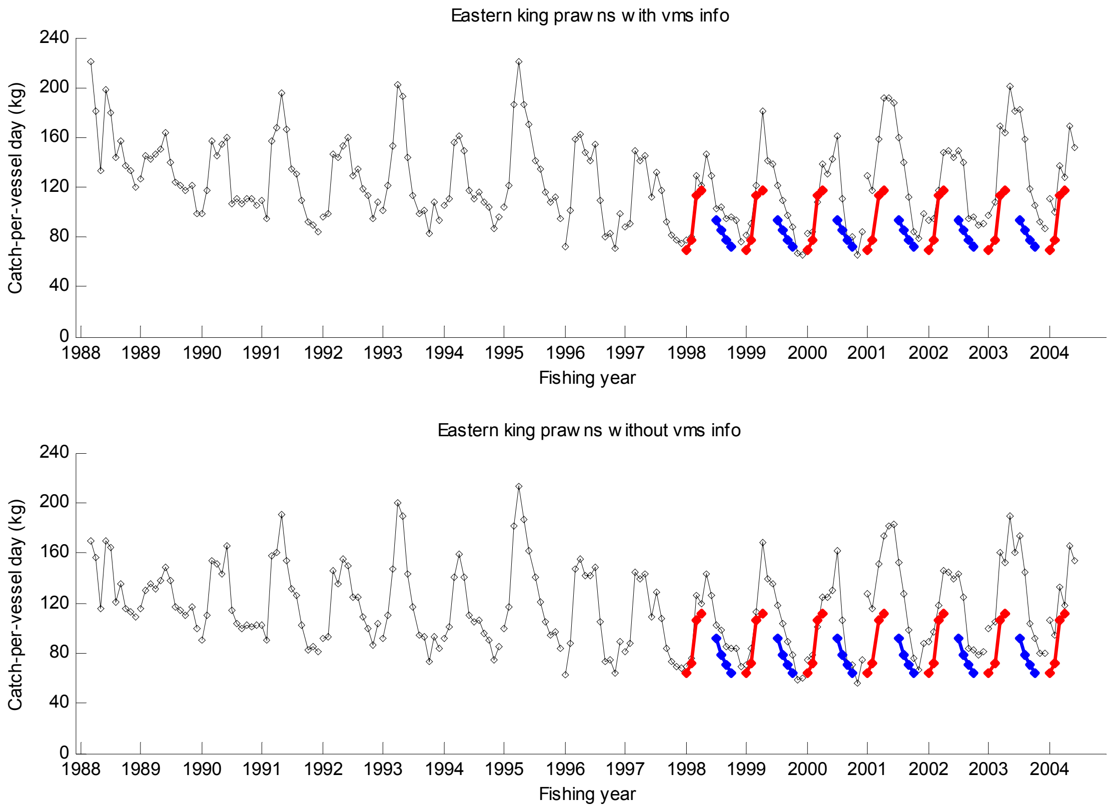


Figure 10.6. Average monthly standardised catch rates predicted from Model A (with VMS information, upper plot) and Model B (without VMS information, lower plot).

10.4 Discussion

The fishery for EKPs is geographically complex and highly seasonal. The southern trawl closure during October restricts fishing to waters deeper than 50 fathoms. Fundamentally different gear configurations are used when fishing for prawns at different depths. In depths less than 50 fathoms (approx. 90 m) the maximum total combined foot-rope and head-rope lengths are restricted to 88 m (shallow fishery) compared to 184 m in depths greater than 50 fathoms (deepwater fishery) (QECTMP 2004). O'Neill and Leigh (2006) considered these factors in their standardisation model. However, several CFISH grids straddle the 50 fathom depth contour and data reported by CFISH grid have insufficient resolution to determine whether most effort was exerted using deep or shallow gear configurations. For these reasons O'Neill and Leigh (2006) restricted their analysis to the 14 CFISH grids for which catch and effort could more reliably be allocated to deep and shallow gear configurations.

In order to assess the influence of VMS parameters on the existing standardisation of CPUE, our analysis was also confined to the 14 CFISH grids selected by O'Neill and Leigh (2006). The factors most influencing catch rates are very well identified in O'Neill and Leigh's (2006) analysis and, not surprisingly, we found that the extension of their standardisation model to include VMS-derived fine-scale fishing effort and area information resulted in only marginal improvements in the CPUE standardisation. This marginal contribution is considered to be due to the difficulty of incorporating the high precision effort and area information into the relatively coarse spatial scale logbook data. Extending the standardisation procedure from the 14 selected CFISH grids to all grids where EKP are caught (see Figure 10.7 – Figure 10.9) is expected to result in a greater contribution of VMS-derived parameters to the model. Potential application of VMS data to identify shallow and deep trawling areas for the EKP fishery is described in Appendix E.

Our use of the 2001–2004 period (for which VMS data are available) to estimate effort and area fished for 1988-2000 may be improved with the inclusion of further years in the model. However, the resolution of the historic data cannot be improved, nor can the estimates resulting from our back calculation be validated. Further improvement to the standardisation of CPUE series extending prior to the availability of high-resolution VMS data (from 2001 on) will therefore remain difficult. For CPUE series commencing in 2001 the inclusion of VMS-derived parameters is expected to greatly improve CPUE standardisations.

The choice of Reference Points (in this case 70% of historic CPUE) at which review events are triggered is best determined in relation to the historic productivity of the stock, ideally indicated by quantitative stock assessments. Recommendation of alternative CPUE reference points or validation of the efficacy of the existing 70% reference point is therefore beyond the scope of the current study. Here we consider how the availability of VMS information may improve the current reference points and allow the development of alternative reference points.

For the Queensland East Coast Trawl Fishery review events are currently defined with reference to catch rates of targeted species in a reference period of 1988 to 1997. These reference years fall within the period for which VMS-derived parameters had to be estimated in our standardisation model. With the extension of the time series of high-resolution VMS data it would be prudent to revise the reference period to include years for which the logbook data may be better validated by VMS information.

The availability of VMS information may also allow the development of alternative measures of the performance of fisheries. For example, the expansion or contraction of the area fished or changes in the concentration of effort are readily measured using VMS information and may be related to changes in the spatial coverage and productivity of the stock. Recommendation of particular VMS-related performance measures (and the choice of appropriate thresholds) requires a relationship between these measures and the productivity of the stock. Estimates of productivity during potential reference periods would be best provided by a spatially explicit stock assessment.

As noted previously, the inclusion of the VMS data resulted in only marginal improvement or changes in the fit of the standardisation model to the catch data. The most likely reason, as identified in the previous chapters, is that catch data are not available at the same scale as the fine-scale effort data. The averaging of catch across small spatial units (e.g. 1 × 1 nm^2 grids) will in effect blur the message in the fine-scale effort data. The need for fine-scale catch data to match the VMS effort data must be emphasised as the factor that would give greatest improvement to stock assessment and the evaluation of the reference points using VMS data.

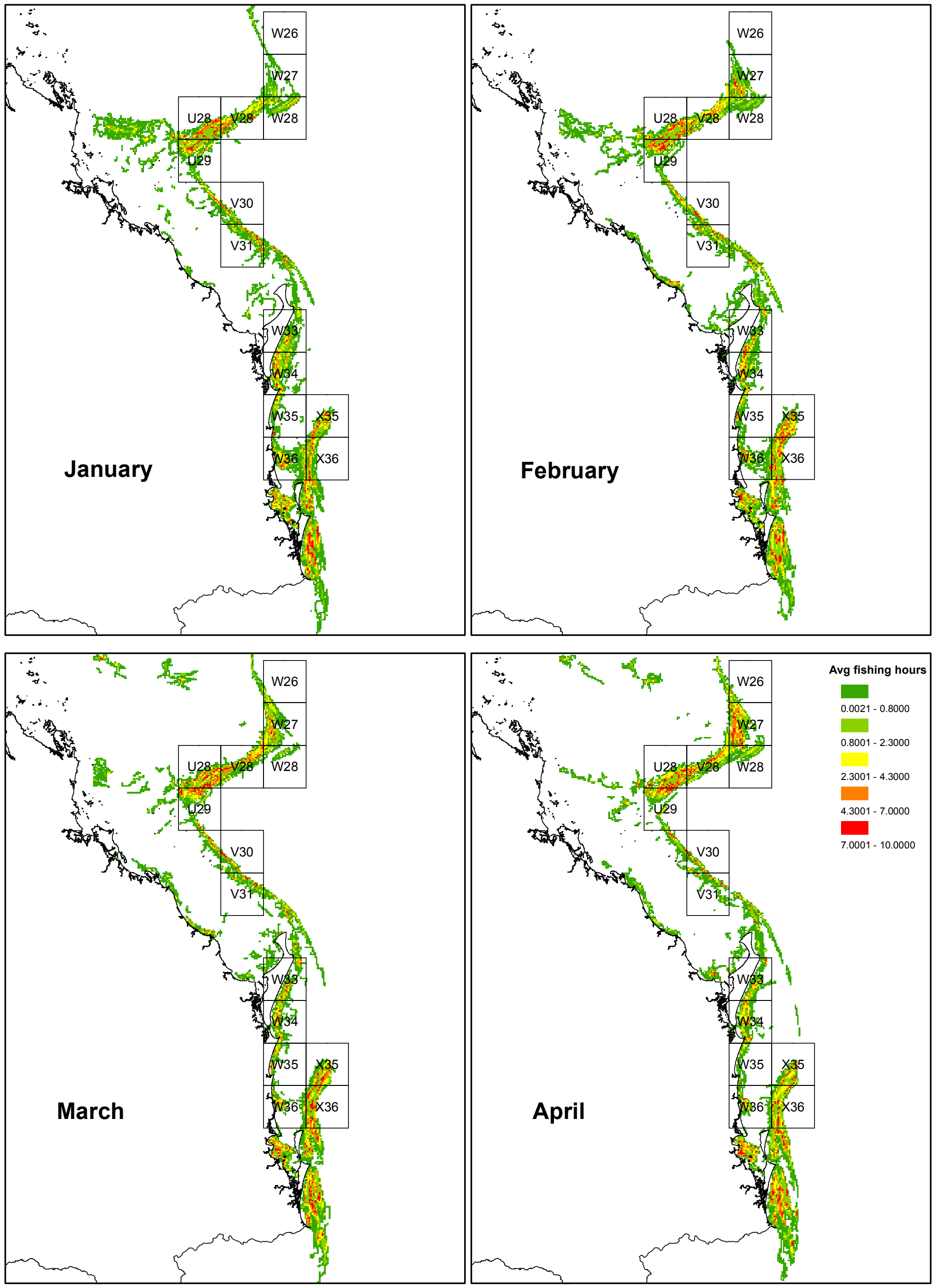


Figure 10.7. Spatial distribution of average monthly fishing effort (hours fished) of the Queensland EKP fishery between 2001 and 2004 (January–April). A spatial resolution of one-minute cell was used to extract the VMS records. The squares are the 30-minute CFISH grids used in the standardisation model.

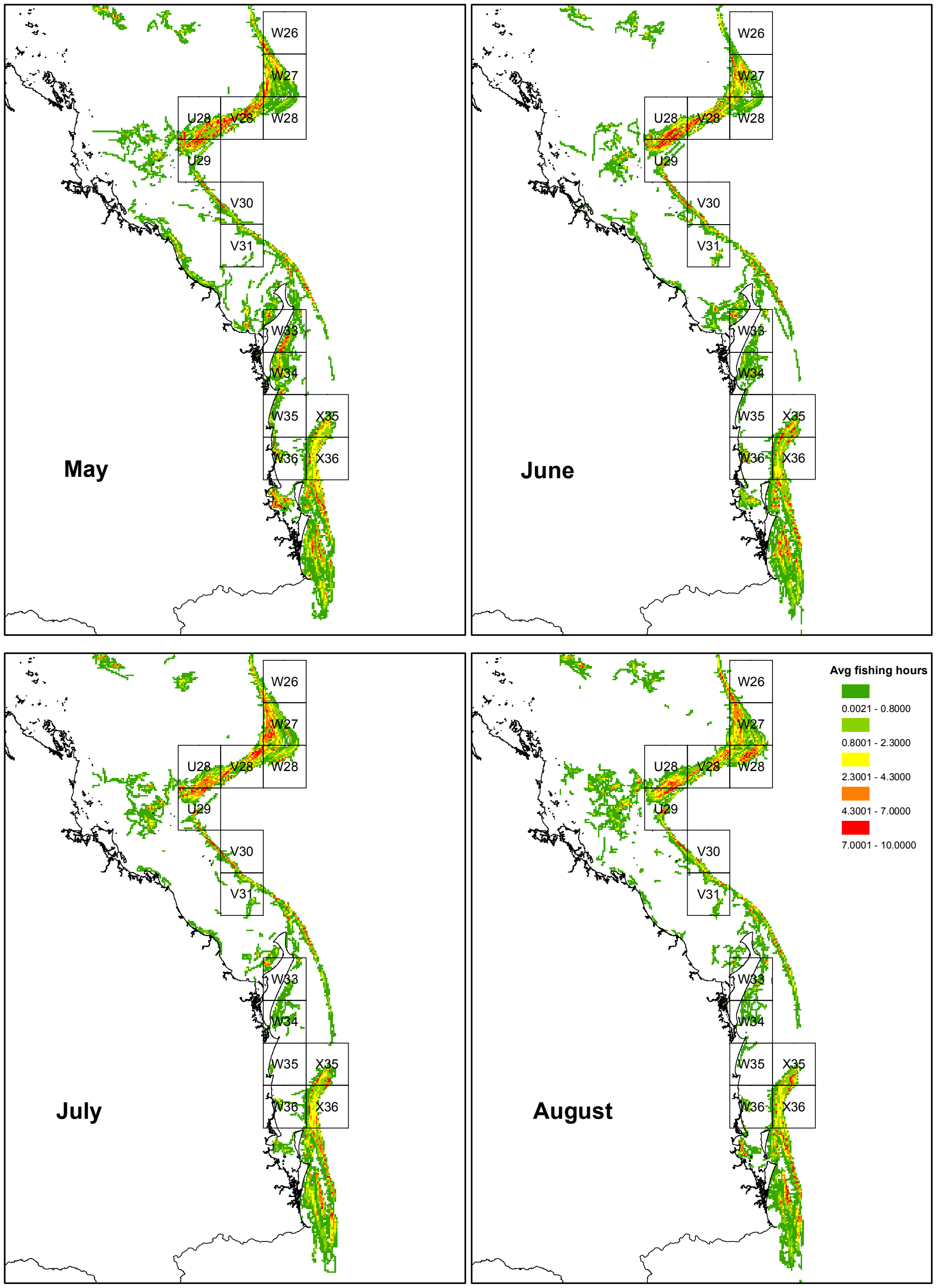


Figure 10.8. Spatial distribution of average monthly fishing effort (hours fished) of the Queensland EKP fishery between 2001 and 2004 (May–August). A spatial resolution of one-minute cell was used to extract the VMS records. The squares are the 30-minute CFISH grids used in the standardisation model.

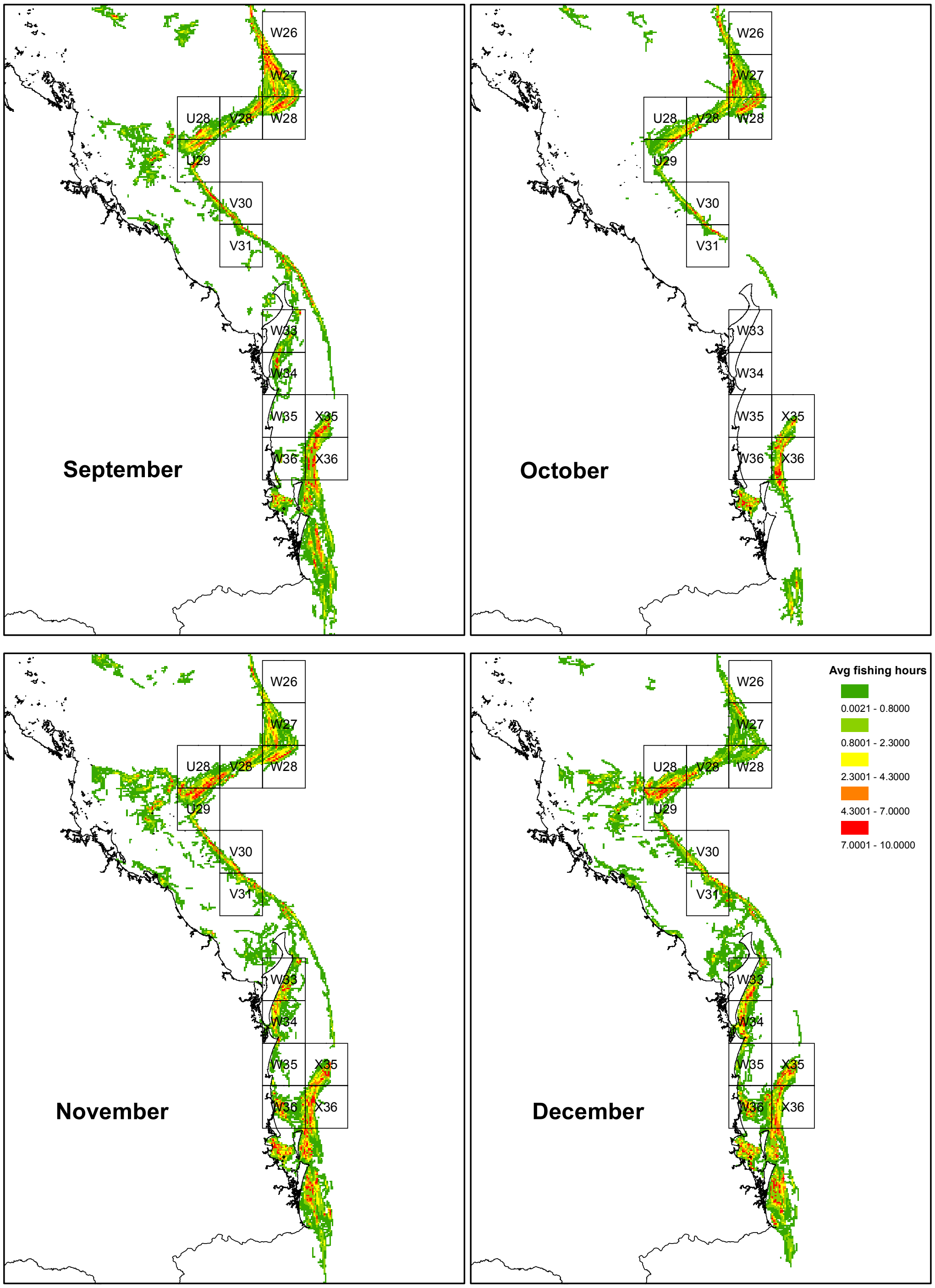


Figure 10.9. Spatial distribution of average monthly fishing effort (hours fished) of the Queensland EKP fishery between 2001 and 2004 (September–December). A spatial resolution of one-minute cell was used to extract the VMS records. The squares are the 30-minute CFISH grids used in the standardisation model.

11 ROLE OF TARGETING AND DEPLETION OF AGGREGATION IN THE QUEENSLAND MAJOR TRAWL FISHERIES

Mai Tanimoto and Rick Officer

11.1 Introduction

Commercial CPUE data are commonly used as indicators of stock abundance in fisheries stock assessment. Relating a CPUE index back to the abundance of the underlying population requires a constant of proportionality associated with the efficiency of the fishing gear. This constant is usually denoted as the 'catchability coefficient', q, and conventionally used in the catch equation:

$$C = qEB$$

Equation 11.1

where C denotes catch, E denotes fishing effort, and B denotes biomass. From this equation CPUE can be defined as:

$$C/E = qB$$

Equation 11.2

In earlier chapters of this report we have demonstrated that VMS data can improve identification of periods of effective effort. Improved resolution in catch reporting would also help to resolve the left side of Equation 11.2. Improved accuracy in catch and effort reporting would therefore improve the accuracy with which biomass can be estimated, provided that catchability is also well estimated.

The relationship between CPUE and abundance can be biased by concentration in the spatial distribution of fishing effort, and by changes in the area fished (Campbell 2004). If such issues could be overcome then it would be theoretically possible to estimate biomass directly when catch and effort are known. However, a contextual distinction in the meaning of q creates an additional complication.

When CPUE is derived from commercial catch and effort data, q is interpreted as the proportion of the population biomass caught by one unit of fishing effort (Francis, Hurst *et al.* 2003). When catch rates refer to trawl survey indices the interpretation of q is slightly different. Here q refers to the product of the survey area and the proportion of the biomass that is caught per unit area swept (Francis, Hurst *et al.* 2003). This distinction has led to a dichotomy in catchability research involving either the investigation of the relationship between population size and effort, or a more direct investigation of fishing gear efficiency (Arreguin-Sanchez 1996). Photography, underwater video and diver observation have been used to investigate gear efficiency directly. Indirect methods such as depletion experiments, mark-recapture studies and change-in-ratio experiments have been the domain of investigation of the relationship between population size and effort

(Gedamke, DuPaul *et al.* 2004). Although depletion experiments have been used in a number of applications (Beukers-Stewart, Jenkins *et al.* 2001; Burridge, Pitcher *et al.* 2006; Gonzalez-Yanez, Millan *et al.* 2006; Joll and Penn 1990; Kangas, Sporer *et al.* 2005; Ralston and Tagami 1992; Wright, Caputi *et al.* 2006), the demanding assumptions of classical depletion analyses have restricted opportunities to use fishery-dependent data. However, the improved spatial and temporal resolution of VMS is providing new opportunities to use fishery-dependent data in depletion analyses (e.g. Deng, Dichmont *et al.* 2005; Gedamke, DuPaul *et al.* 2004; Harrington, Semmens *et al.* 2006).

In this chapter we consider whether VMS data can be used to improve the estimation of q. We first explore the potential of using VMS data to identify spatial concentrations in fishing effort and the targeting of areas of high abundance. Secondly we consider whether high-resolution VMS data can be used to estimate catchability directly. Our objective is to evaluate the capacity of VMS data to improve the quality and credibility of future stock assessment.

11.2 Method: Analysis of spatial concentration of effort

These analyses were conducted for the EKP, tiger/endeavour prawns, and scallop fisheries. Fine-spatial-scale annual effort data of these three fisheries were obtained by using TrackMapper for the 2001–2005 fishing years. The fishing year was defined from November through to October inclusive for EKP and scallop and as the calendar year for tiger/endeavour prawns (O'Neill and Leigh 2006). Because VMS data were not available before December 2000, the annual effort of EKP and scallop fisheries in the 2001 fishing year was obtained from December 2000 to October 2001 (11 months of records).

The area covered by the fishery and the total amount of effort have remained relatively stable for the EKP fishery (Figure 11.1). Close examination of the catch and effort in each one-minute cell for each fishery shows that many cells received very little catch or effort. Scallops, EKP and tiger/endeavour prawns caught in these cells are unlikely to have been targeted. The purpose of this analysis was to investigate the targeting behaviour in these fisheries. We therefore selected only those cells which accounted for the top 99% of total annual catch for each year and fishery. Eliminating 1% of the total catch resulted in a substantial reduction in the number of cells considered in further analyses (approximately a 50% reduction for the prawn fisheries, and a 60% reduction for the scallop fishery).

The extent to which fishing effort was targeted towards aggregations was analysed by plotting cumulative fishing effort against area fished after sorting the one-minute cells by a) decreasing effort and b) decreasing CPUE (Campbell 2004). These plots indicate whether there was a spatial concentration in the fishing effort, and whether the distribution of effort favoured higher CPUE cells (i.e. the targeting of aggregations). We have also added a plot of cells fished by decreasing catch to provide a more complete picture of the fisheries.

11.3 Results: Spatial concentration of fishing effort

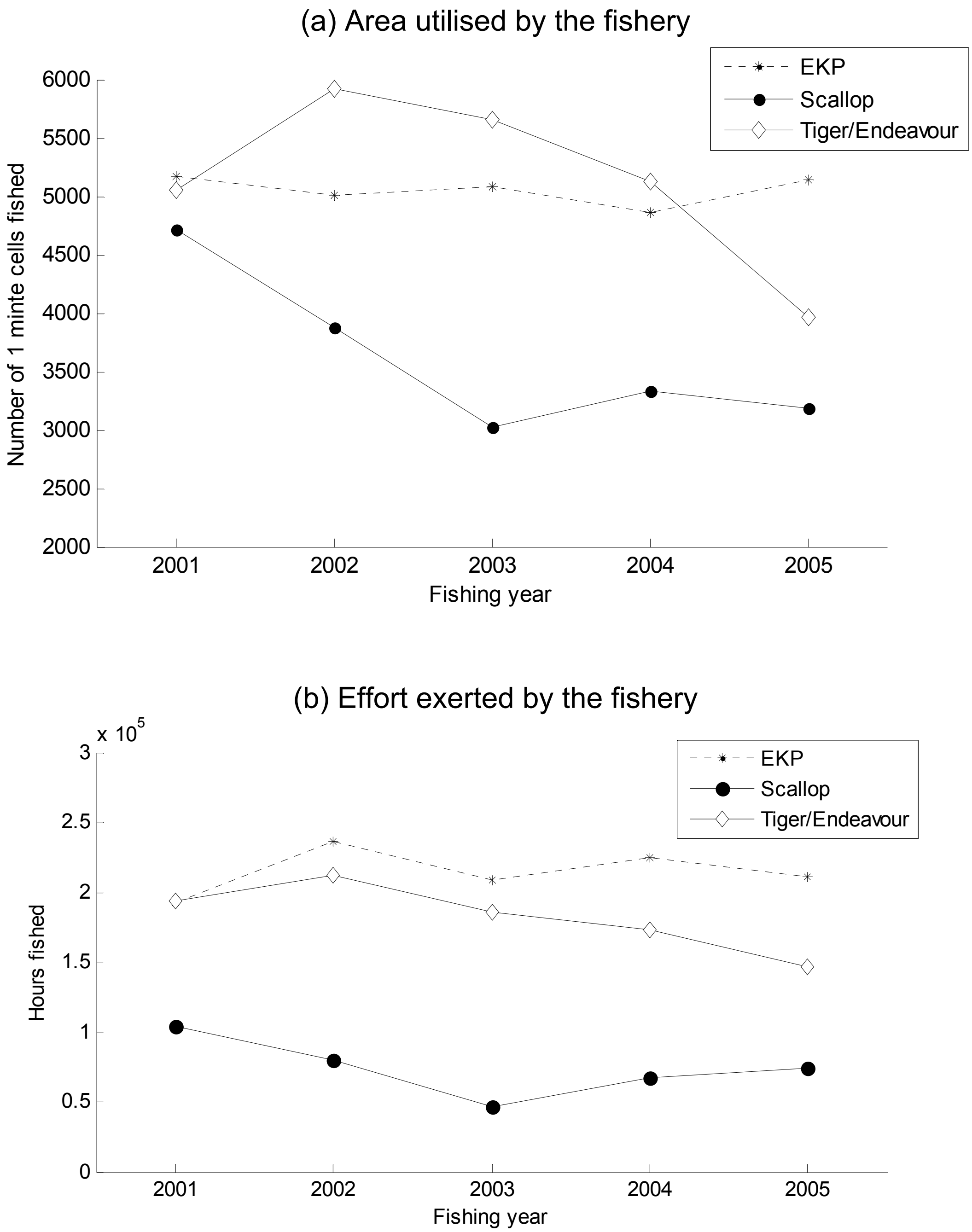


Figure 11.1. Temporal trends in (a) the distribution of fishing effort, and (b) the total number of hours fished in the EKP, scallop and tiger/endeavour prawn fisheries. The time series shown from 2001–2005 is the period for which VMS data are available.

11.3.1 *EKP fishery*

Spatial concentration of fishing effort

For the EKP fishery, about 80% of the effort was expended in the 60% of the one-minute cells fished and 50% in the 30% of the one-minute cells, indicating that the distribution of the fishing effort is relatively aggregated across the fishing areas (Figure 11.2 (a)). In general, the degree to which fishing effort was concentrated seemed to be consistent across fishing years, indicating the relatively constant temporal distribution of fishing effort.

Targeting of aggregations

About 50% of the effort was spent in the top 45–60% of the high CPUE cells and 20% in the top 23–30% of cells. This indicated that there was no clear evidence of effective targeting of the areas with higher CPUE (Figure 11.2 (c)). The degree of the spatial targeting of fishing effort seemed to vary across fishing years with the least effective targeting in the 2001 fishing year and most effective in 2002.

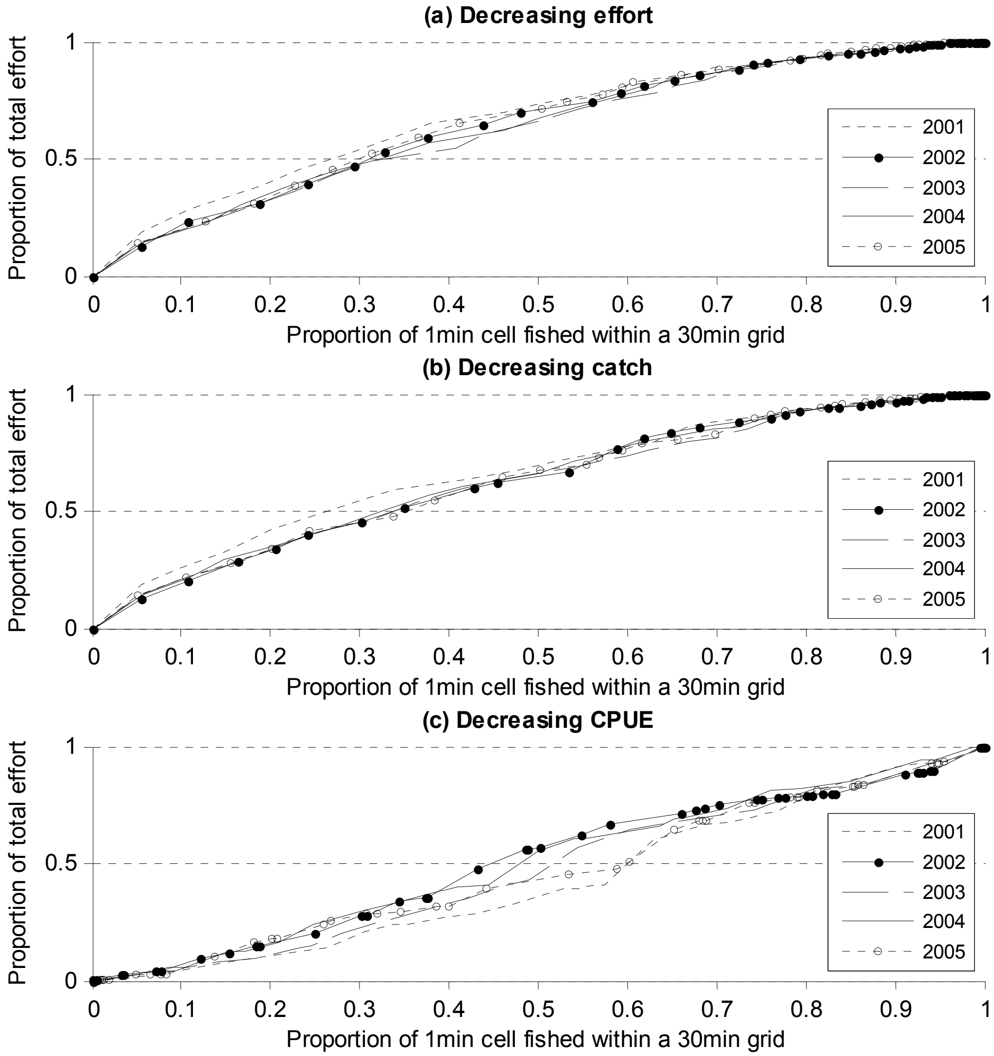


Figure 11.2. Cumulative effort vs. cumulative area fished for the EKP fishery between fishing years 2001 and 2005, after ordering one-minute cells by (a) decreasing effort, (b) decreasing catch, and (c) decreasing CPUE. Both effort and area fished were expressed as a proportion of the respective annual totals.

11.3.2 *Scallop fishery*

Spatial concentration of fishing effort

For the scallop fishery, about 80% of the effort was expended in the 58–70% of one-minute cells fished and 50% in the 25–38% of one-minute cells, indicating the existence of the spatial aggregation in the fishing effort (Figure 11.3 (a)). The degree of the spatial aggregation of fishing effort slightly varied across fishing years with the least aggregated effort in the 2001 fishing year and the most aggregated effort in 2003.

Targeting of aggregations

In general, about 50% of the effort was spent in the top 45–60% of the high CPUE cells and 20% in the top 23–30% of cells. This indicated that there was no clear evident of effective targeting of the areas with higher CPUE (Figure 11.3 (c)). The degree to which fishing effort was spatially targeted varied between fishing years and was least effective in the 2005 fishing year and most effective in 2002. It is important to note that there were a number of scallop replenishment areas implemented between the fishing years 2001 and 2005, and these closures may have contributed to the different patterns of the spatial targeting of fishing effort among the fishing years.

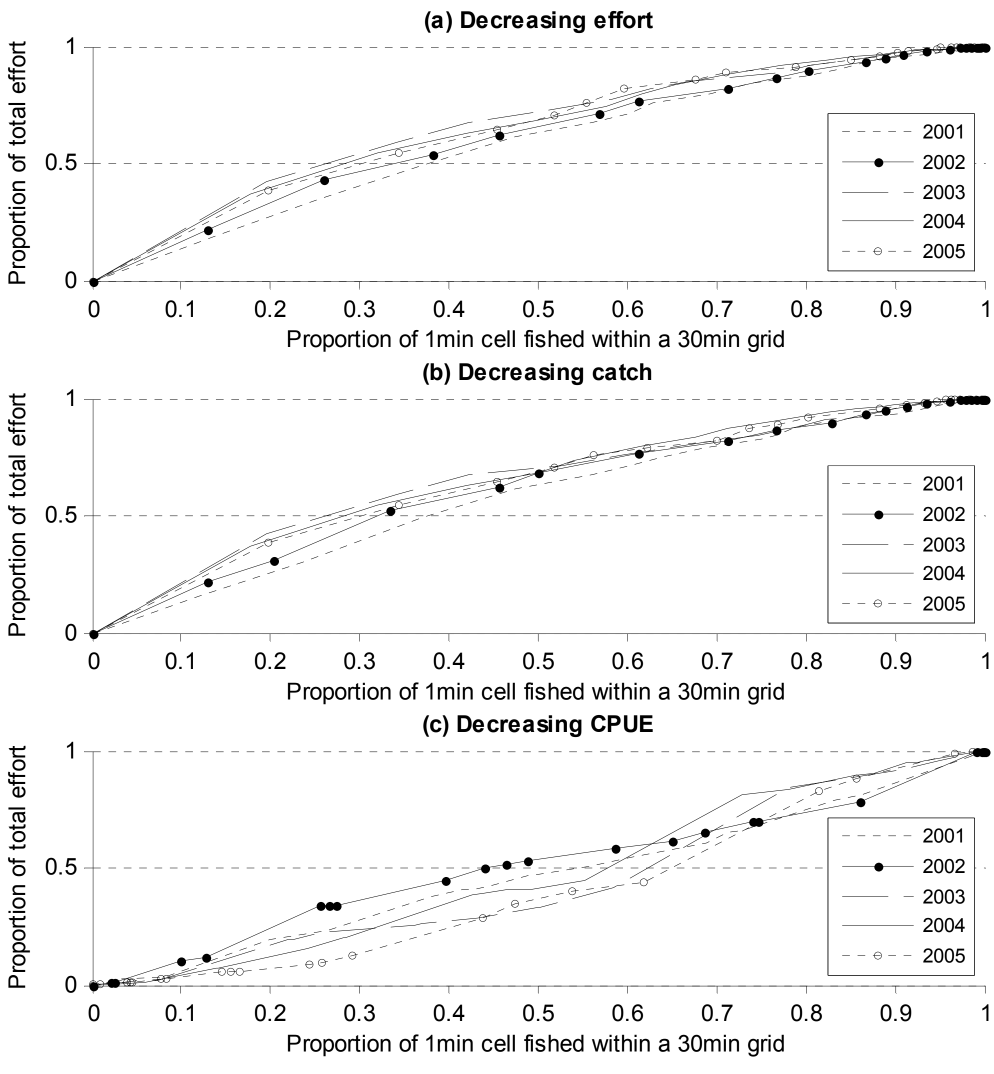


Figure 11.3. Cumulative effort vs. cumulative area fished for the scallop fishery between fishing years 2001 and 2005, after ordering one-minute cells by (a) decreasing effort, (b) decreasing catch, and (c) decreasing CPUE. Both effort and area fished were expressed as a proportion of the respective annual totals.

11.3.3 *Tiger/Endeavour prawns fishery*

Spatial concentration of fishing effort

For the tiger/endeavour prawns fishery, about 60% of the effort was expended in the 40% of one-minute cells fished and 20% in the 8–10% of one-minute cells, indicating that the distribution of the fishing effort is relatively aggregated across the fishing areas (Figure 11.4 (a)). In general, the degree of the aggregation of fishing effort was fairly consistent across fishing years, indicating the constant temporal distribution of fishing effort.

Targeting of aggregations

About 60% of the effort was spent in the top 40–50% of the high CPUE cells and 30% in the top 20–30% of cells, indicating that the fishing effort was relatively concentrated in those cells with a higher CPUE (Figure 11.4 (c)). The degree to which fishing effort was spatially targeted was relatively consistent across the fishing years.

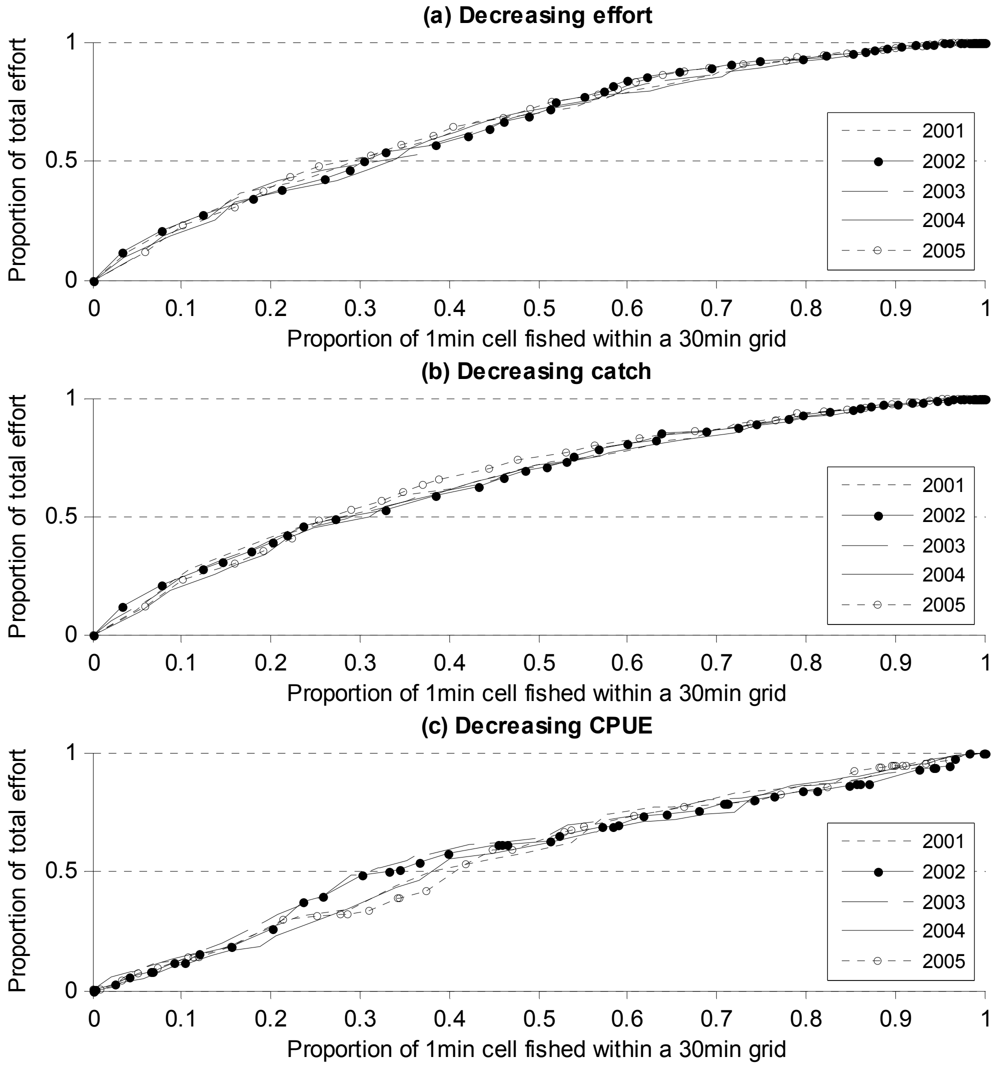


Figure 11.4. Cumulative effort vs. cumulative area fished for the tiger/endeavour prawns fishery between fishing years 2001 and 2005, after ordering one-minute cells by (a) decreasing effort, (b) decreasing catch, and (c) decreasing CPUE. Both effort and area fished were expressed as a proportion of the respective annual totals.

11.4 Method: Depletion analysis

The sedentary characteristics of scallops and the management arrangements of the scallop fishery (i.e. closed fishing areas) provide suitable circumstances for the application of depletion analyses. A DeLury depletion analysis was carried out on the data obtained from TrackMapper to investigate the catchability and gear efficiency of the Queensland scallop fishery.

Depletion analysis was not conducted for the EKP fishery due to a) no major closure, b) its highly migratory behaviour, and c) the strong lunar phase effect on catches. The tiger/endeavour prawns fisheries were originally considered, but it was found that there was too limited temporal and spatial distribution of successive fishing effort to investigate the depletion of the stock.

11.4.1 *Description of depletion analysis*

The DeLury depletion model (DeLury 1947) is frequently used to estimate the catchability coefficient q and initial population number N_0 from catch and effort data (Sanders 1988). A linear relationship between the logarithm of CPUE and cumulative effort can be modelled as:

$$\log_e(CPUE_t) = \log_e(q \cdot N_0) - q \cdot cumeff_t$$

Equation 11.3

where $CPUE_t$ = the catch rate at time t, N_0 = initial abundance, $cumeff_t$ = cumulative effort at time t and q = the catchability coefficient.

The catchability coefficient is generally defined as the proportion of the population taken by one unit of fishing effort (Gedamke, DuPaul *et al.* 2004; Ricker 1975; Swain and Sinclair 1994). The catchability coefficient is more specifically defined by Paloheimo and Dickie (1964) as:

$$q = \frac{Ea}{A}$$

Equation 11.4

where E = gear efficiency (e.g. the proportion of fish in the trawl area that will be caught by the fishing gear), and a = the area trawled by the gear in one unit of fishing effort. A is defined either as the total area of the population (Paloheimo and Dickie 1964), or as a fishing ground of size A (Gulland 1983) or the area of the study site (Gedamke, DuPaul *et al.* 2004). The former definition of A assumes that the area trawled by one unit of effort a is representative of the entire area of the fishery. This assumption may be difficult to satisfy in populations that are not uniformly distributed over the stock area (Swain and Sinclair 1994), or that occur on grounds which variably impact on the performance of the fishing gear. The definition of the study area used here is most akin to that of Gedamke (2004).
DeLury analyses require that the following assumptions are satisfied:

1) the population is closed (the effects of migration and natural mortality are negligible)
2) constant catchability during the study period for each individual (random spatial distribution of fishing effort)
3) fish removal must significantly reduce the population size, and
4) all of the expended fishing effort is accounted for (Cowx 1983; Gedamke, DuPaul *et al.* 2004; Ricker 1975; Seber 1982).

The selection of study areas which satisfy these assumptions is therefore crucial to successful depletion experiments.

11.4.2 *Selection of study site and period of study*

The VMS database was used to select study areas that best satisfied the assumptions of DeLury analyses. Queensland's scallop fishery has several characteristics that make it suitable for depletion analyses. Since 1997 the fishery has largely operated within a series of spatial closures (scallop replenishment areas, SRAs). These SRAs have been periodically opened to pulse fishing on a rotating basis with slight changes in the configuration (Figure 11.10–Figure 11.12 and Table 11.5–Table 11.7). Catch and effort are most substantial within the SRAs during the first 1 to 2 weeks following opening (see Figure 11.5). Most scallops are taken from the fishing ground over a short period of time, satisfying the assumption of DeLury analyses that fish removal must result in a significant reduction of the population size. The requirement that emigration, immigration and natural mortality remain negligible was assumed to be satisfied over the 1 to 2 week periods of fishing following each opening. Six SRAs were selected as study sites with different study periods based on their own temporal distributions of catch and effort (
Table 11.1).

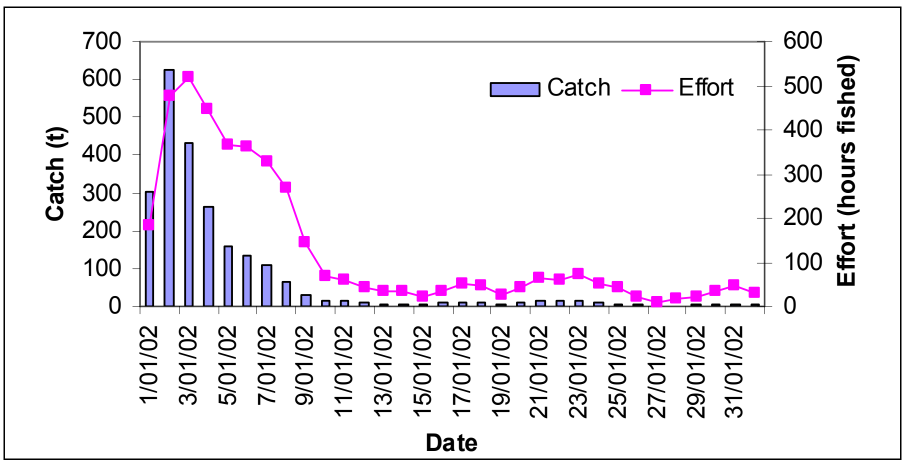


Figure 11.5. Example of the temporal distribution of catch and effort after the opening of scallop replenishment areas (SRAs). This figure shows total daily catch and effort taken from the Yeppoon B SRA (YPB) for the first 31 days after the opening on 1 Jan 2003.

Table 11.1. The selected Scallop Replenishment Areas (SRAs) and opening dates which observed high catch and effort records.

SRA	Opening Date	Study period (days)	Data name
Bustard Head	1-Feb-01	16 (1 - 16-Jan-01 inclusive)	BH01
Yeppoon Area B	1-Jan-02	8	YB02
Bustard Head Area B[1]	1-Jan-03	6	BHB03
Bustard Head Area B[2]	1-Jan-04	18	BHB04
Bustard Head Area A[2]	1-Jan-05	8	BHA05
Hervey Bay Area A	1-Jan-05	8	HBA05

[1]: Configuration at the implementation of the *Fisheries (East Coast Trawl) Management Plan 1999* (Figure 11.11),
[2]: Current SRA configuration (Figure 11.12).

11.4.3 *Data extraction*

In order to meet the assumption of random spatial distribution of fishing effort, we needed to disaggregate catch and effort data into finer spatial strata (Hilborn and Walters 1992). For this study TrackMapper was used to extract catch and effort records at the fine spatial scale of a ¼ × ¼ minute, referred to as a ‘cell’ hereafter. This spatial resolution is 576 times finer than that of the current logbook recording system (6 × 6-minute grid). On such a fine scale, fishing effort appeared to be deployed in a random spatial manner (see Figure 11.6).

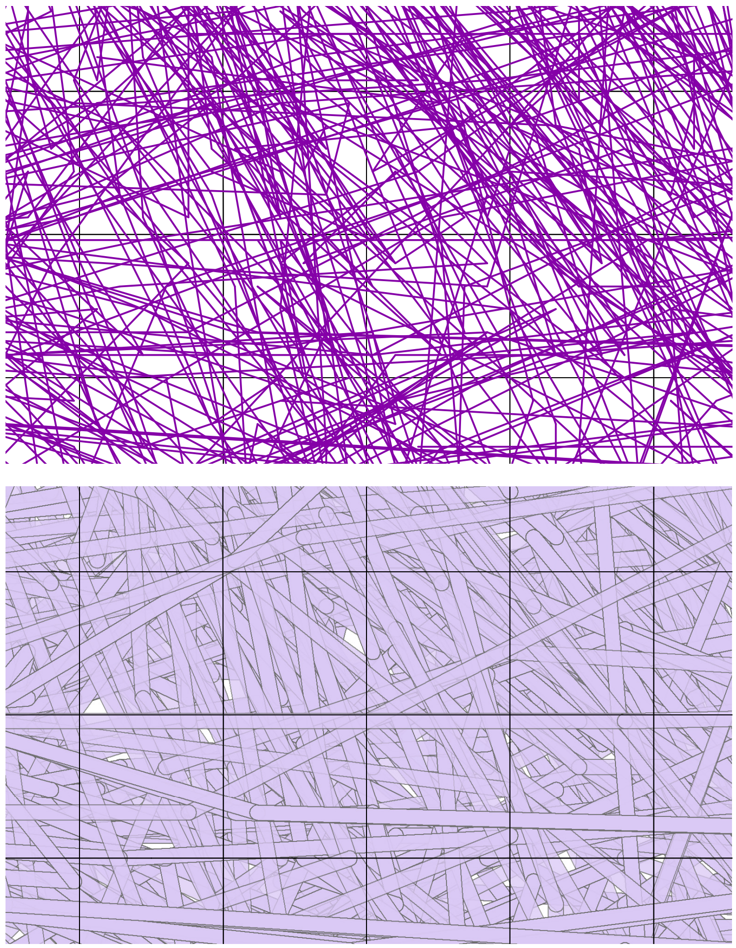


Figure 11.6. The VMS trawling lines (top) and the simulated trawling tracks (bottom) in one of the study areas. The ArcView BUFFER function was used to generate the trawling tracks with width of 34.6 m, the standard tow width in the scallop fishery (O'Neill and Leigh 2006). Grid lines delimit ¼ × ¼ minute cells.

11.4.4 *Cell selection*

Within each of the selected study areas, not all cells will satisfy the assumptions of depletion analysis. Further filtration of the catch and effort data was required to ensure that sufficient fishing effort was exerted over the study period. We selected cells in which:

1) at least 3 days of non-zero catch and effort records was recorded over the study period; and,
2) whose total fishing effort amounted to the top 10–90% of the overall fishing effort for the entire study area over the study period.

The first selection criterion provides a minimum number of points through which to fit the depletion curve. The second criterion was applied after data with less than two days of records were eliminated. This criterion aimed to ensure that sufficient effort was exerted to deplete the stock within each cell. The influence of the second selection criterion was examined using sensitivity analyses which applied a total of six levels of selection settings for each dataset (Table 11.2). For example, Run B included the top 90% of effort data which had at least three days of records over the study period. Data selection ranged from least restrictive in Run A (the first selection criteria only) to most restrictive in Run F.

Table 11.2. The six levels of data restriction applied as criteria for cell selection.

Run	Minimum number of days	Cut-off value of total fishing effort in a cell	Description
A	3	No cut-off value	Include all data with >3 days of record
B	3	10th percentile	Exclude lowest 10% of the effort data (include top 90% of effort data)
C	3	25th percentile (lower quartile)	Exclude lowest 25% of the effort data (include top 75% of effort data)
D	3	50th percentile (median)	Exclude lowest half of the effort data (include top 50% of effort data)
E	3	75th percentile (upper quartile)	Exclude lowest 75% of the effort data (include top 25% of effort data)
F	3	90th percentile	Exclude lowest 90% of the effort data (include top 10% of effort data)

11.4.5 *Statistical analysis*

The catchability coefficient was estimated by fitting a general linear regression shown as:

$$\log_e(CPUE_{i,t}) = cell_i + cell \cdot cumeff_{i,t}$$

Equation 11.5

where $CPUE_{i,t}$ is the catch rate in the i^{th} cell at time t, $cell_i$ is the effect of the i^{th} cell and $cumeff_{i,t}$ is the cumulative effort for the i^{th} cell at time t. By including the effect of cell as a factor, and by adding the interaction term between cell and cumulative effort ($cell.cumeff_{i,t}$), individual catchability coefficients were estimated for each cell in one model (rather than by fitting depletion curves for each cell individually). The model therefore has more residual degrees of freedom, giving more power to the test of significance of the catchability coefficient estimates (hence reducing Type II error). The parameter estimates of the $cell$ term and the interaction term $cell \cdot cumeff_{i,t}$ are equivalent to the constant term $\log_e(q \cdot N_0)$ and the slope of -q respectively, in Equation 11.3. The statistical software package GenStat Version 8 (GENSTAT 2005) was used to undertake the analysis. Once regression analyses were conducted for all study areas and sensitivity runs, the significance of catchability coefficient estimates was examined for each cell. Summary statistics of ***significantly positive q*** were presented for each sensitivity run and study area. This means that if the catchability coefficient estimate in a cell *X*, for example, was not statistically different from zero or negative (i.e. had a positive slope), catch and effort data in that cell were not considered suitable for the depletion analysis and were therefore eliminated from the results.

Once catchability coefficients were identified for the respective cells, gear efficiency was estimated for each area by applying the median of respective q with the relationship:

$$E = q\frac{A}{a} = q\frac{\text{area of 1/4nm} \times \text{1/4nm grid cell}}{\text{tow speed} \times \text{tow duration} \times \text{gear width}}$$

Equation 11.6

The standard tow speed and total gear width for the Queensland scallop fishery were set to be 2.5 knots and 34.6 m (0.0187 nm) respectively (O'Neill and Leigh 2006). One unit of trawl duration was set to equal one hour of trawling.

11.5 Results: Depletion analysis

The q estimates for each selection criteria and study area are shown in Figure 11.7. When we included all the cells (Run A) or cells with smaller total effort (Run B–Run C), extremely high q estimates were observed. As we selected cells with higher fishing effort, cells with extreme q estimates were eliminated from the analysis and more realistic q coefficients were obtained. The results of depletion analysis for restrictive runs (Run D–Run F) were further detailed in Table 11.3.

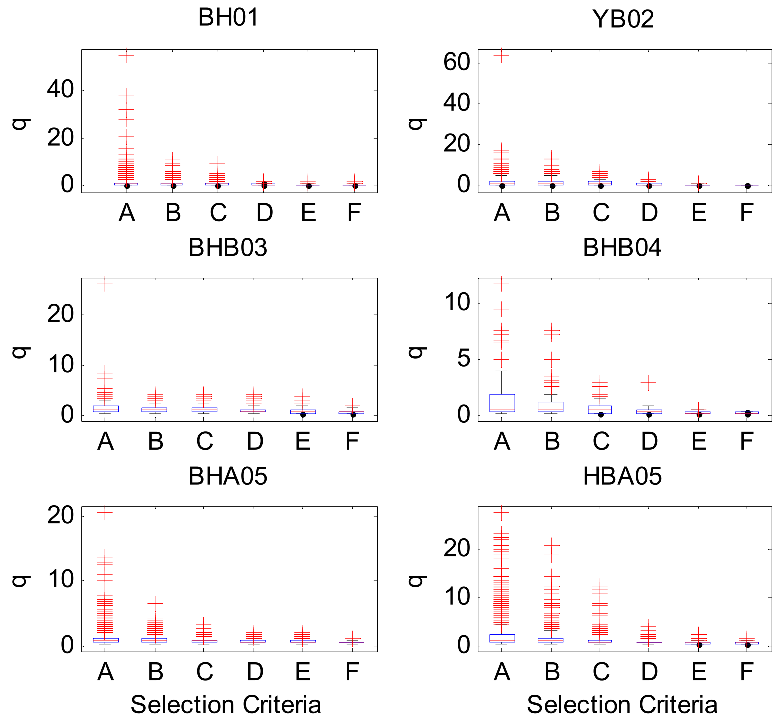


Figure 11.7. Box plots of the estimates of catchability coefficient for each selection criteria and study area.

Table 11.3. Results of estimated catchability coefficients and gear efficiencies for Runs D–F for each area.

BH01

Run	D	E	F
cut-off effort (hrs)	2.78	4.20	5.42
d.f.	3462	1804	735
% significant grid	87	92	96
q (mean)	0.66	0.56	0.51
q (median)	0.62	0.53	0.48
STD (*q*)	0.24	0.22	0.24
E	0.83	0.71	0.64

YB02

Run	D	E	F
cut-off effort (hrs)	1.33	2.62	4.34
d.f.	3426	1946	809
% significant grid	37	57	69
q (mean)	0.81	0.53	0.41
q (median)	0.66	0.48	0.38
STD (*q*)	0.49	0.23	0.12
E	0.88	0.64	0.50

BHB03

Run	D	E	F
cut-off effort (hrs)	2.02	2.93	3.85
d.f.	762	404	168
% significant grid	29	34	42
q (mean)	1.00	0.85	0.69
q (median)	0.80	0.60	0.57
STD (*q*)	0.70	0.66	0.39
E	1.07	0.80	0.76

BHB04

Run	D	E	F
cut-off effort (hrs)	1.95	3.28	4.51
d.f.	3933	2354	1030
% significant grid	6	9	11
q (mean)	0.44	0.31	0.25
q (median)	0.35	0.26	0.22
STD (*q*)	0.44	0.12	0.07
E	0.47	0.35	0.29

BHA05

Run	D	E	F
cut-off effort (hrs)	3.08	4.41	5.49
d.f.	2471	1326	549
% significant grid	79	86	92
q (mean)	0.65	0.60	0.55
q (median)	0.60	0.53	0.49
STD (*q*)	0.25	0.26	0.24
E	0.80	0.71	0.66

HBA05

Run	D	E	F
cut-off effort (hrs)	2.54	4.38	5.91
d.f.	2257	1213	499
% significant grid	58	74	81
q (mean)	0.68	0.54	0.50
q (median)	0.60	0.49	0.44
STD (*q*)	0.39	0.27	0.24
E	0.81	0.65	0.59

* = percentage of cells with significantly positive *q* ($p < 0.05$) out of all cells in each area and run.

The median of catchability coefficient estimated from the depletion analyses ranged from 0.22 (BHB04, Run F) to 1.07 (BHB03, Run D) which produced efficiencies ranging from 29% to >100% respectively. Standard deviations of q are high for all runs in each area, indicating that *q* estimates vary from cell to cell even in the restrictive runs (see Table 11.3 and Figure 11.9). Unrealistically high efficiency estimates (E >1.0) were obtained for all runs in BHA02, and relatively small estimates were obtained in BHB04. The efficiency estimates obtained from other areas were generally similar to each other, which ranged from on average 63% in Run F to 88% in Run D.

For all study areas, the lowest catchability coefficient (and therefore the lowest gear efficiency) was obtained from Run F which only included cells in the top 10% of total fishing effort in respective study areas and study periods. The sensitivity analysis indicated that the model was quite sensitive to the selection of study cells, and the level of total effort spent in a cell was a critical factor in a depletion analysis.

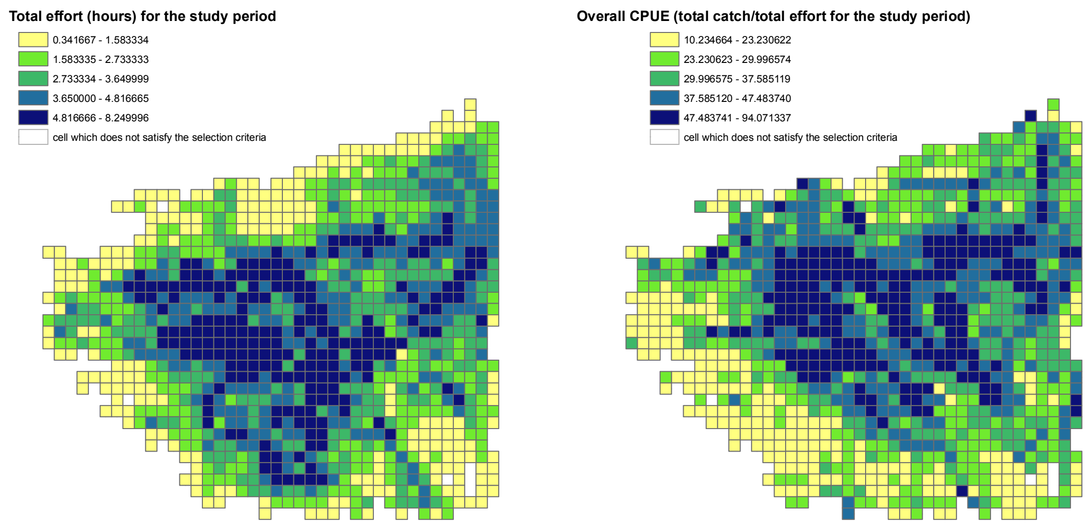


Figure 11.8. The maps of total fishing effort (left) and CPUE (right) from one of the study areas obtained from TrackMapper. The highest total effort was observed in the central and south-central parts of the area, whilst the highest CPUE was observed slightly north of these areas.

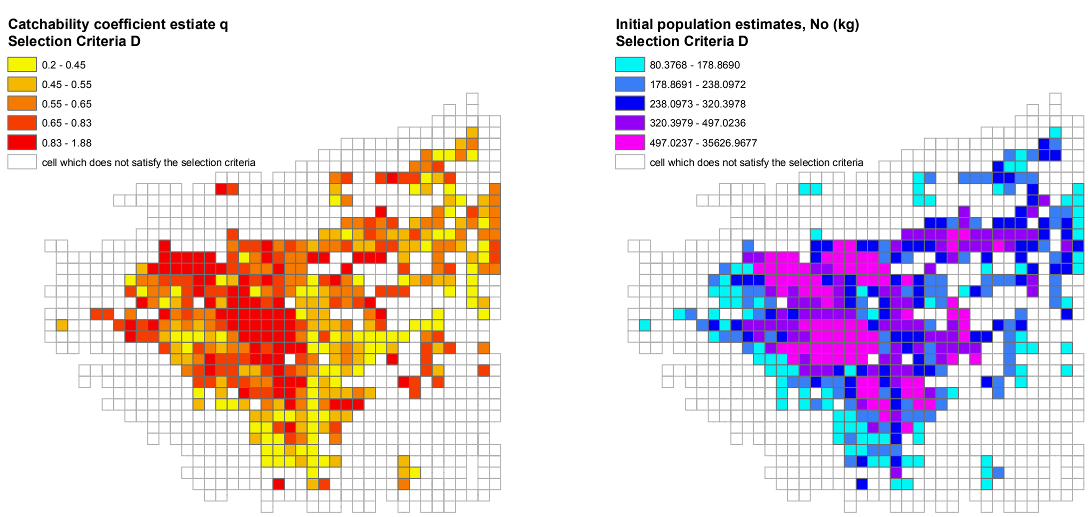


Figure 11.9. The maps of q estimates (left) and initial population (N_0) estimates (right) in one of the study areas which satisfied the selection criteria D. The range of qs was estimated from the analysis, and high q estimates tended to be obtained around the left centre of the area. Similar patterns were found in the distribution of initial population (N_0).

11.6 Discussion

11.6.1 *Targeting*

The results showed that fishing effort for all three fisheries – EKP, scallop and tiger/endeavour prawns – was not randomly distributed but focused on discrete areas. The degree of spatial concentration in the fishing effort was generally homogenous over time for all three fisheries.

The targeting of a greater proportion of the fishing effort towards areas with higher CPUE occurred in the tiger/endeavour prawns fisheries – interpreted as fishers targeting aggregations of prawns. At coarse spatial scale, such concentrations of fishing effort could bias average CPUE high (Campbell 2004). Analysis of the scallop and EKP fisheries, in contrast, showed that although effort was aggregated it was not related to areas of high CPUE – interpreted as fishers following spatial patterns determined by external processes which might include management closures, fuel prices, and cost-benefit decisions.

Although the degree to which effort was targeted varied slightly every fishing year for the EKP and scallop fishery (Figure 11.2 (b) and Figure 11.3 (b)), there was no obvious trend indicating an increasing level of targeting effort over time. This does not necessarily mean that spatial targeting of fishing effort is not a feature of the EKP and scallop fisheries, but it is likely that the 5-year time series of VMS data is too short to detect any trend in the level of effort concentration over time.

These results provide a clear direction for the improvement of future stock assessments. Future calculation of CPUE indices should involve the application of appropriate spatial weightings to correct for the concentration of fishing effort within the fishery area, and the targeting of areas with higher catch rates (Campbell 2004). We also recommend that the spatial concentration of fishing effort continues to be monitored in these fisheries.

11.6.2 *Gear efficiency*

The depletion coefficients estimated from BHB04 were relatively low compared with the other areas. Notably, the cut-off value of total effort was quite small for the 18-day study period, and the proportion of statistically significant cells was less than 11% of the entire cells. These results indicated that there was insufficient fishing effort in this study area to noticeably deplete the population and therefore it was unlikely to be suitable for a depletion analysis. Deng *et al.* (2005) found that the amount of effort in the study area was a crucial factor in depletion analyses. Our results also revealed that q estimates were quite sensitive to the selection of the area (cells) in relation to the amount of fishing effort exerted within the study areas.

The gear efficiency estimates obtained from other study areas were relatively consistent between the same sensitivity runs. However, the results were not strictly comparable between areas due to the different cut-off effort values and study periods applied in each area. The analyses showed that the trawl efficiency for scallops is quite high, ranging from 50% to 76% even from the most restrictive run (Run F). Similar results were found in the depletion experiment conducted by Joll and Penn (1990) for the scallop (*Amusium balloti*) fisheries in Western Australia, which concluded that the commercial otter trawl efficiency was consistently high, with approximately 60% of the scallops in swept area being taken.

11.6.3 *Catchability coefficient*

In general, catchability coefficients estimated from this depletion analysis seemed to be unusually high ($\approx 10^{-1}$) compared with those of Gedamke *et al.* (2004) (i.e. $\approx 10^{-5}$). It is important to note, however, that the magnitude of catchability coefficients is dependent on the unit of fishing effort and the size of A in Equation 11.4. This means that the catchability coefficient is inversely proportional to the area, A (Paloheimo and Dickie 1964). This study used a fine spatial resolution of ¼ x ¼ minute cells ($\approx$ 0.0625 nm^2) with effort units of one hour of trawling, while Gedamke *et al.* (2004) used an analysis cell with radius of 1.9 nm (= 11.34 nm^2) with effort units of one minute of fishing (i.e larger A with smaller units of effort). Therefore, the high level of q estimated from this study was likely to be due to the finer spatial resolution of A, our study areas.

As Arreguin-Sanchez (1996) noted, the meaning and interpretation of catchability will differ depending on how population units are chosen. This implies that the interpretation of q depends on the definition of the term A in Equation 11.4.

The distinction in the meaning of q when considered in the context of CPUE indices versus trawl survey indices proved to be an important factor in our study. Whilst we used commercial CPUE data, the high resolution provided by VMS meant that we were using these data at a spatial scale which is normally the domain of research survey indices.

Our estimates of q are therefore more akin to estimates of gear efficiency rather than fleet catchability. At this scale, q is a function of the vulnerability and availability of the target species, strictly pertains only to the area swept by the fishing gear, and should not exceed 1 (Francis, Hurst *et al.* 2003). In our study, q exceeded 1 in many cells. Values of q that exceed 1 imply that the action of fishing has somehow increased the availability of fish for capture above 100% of the number initially within the fished area. Whilst this could possibly occur through herding of fish into the swept area, or through the attraction of fish from areas adjacent to the area fished, these explanations are implausible in our study of a sedentary scallop species.

It appears that the algorithm we used for the allocation of catch to effort tracks sometimes over-allocates catch to particular cells (see Section 7.2.5). This effect can be seen in Figure 11.8. Several peripheral cells that received amongst the lowest levels of effort are assigned very high catch rates. This would have occurred when a high catch record was distributed across several cells.

If we have a single and reliable estimate of gear efficiency E and the total area of scallop population A, it is technically possible to estimate a catchability coefficient which represents the entire fishery (i.e. q used in a stock assessment) by using Equation 11.4. The total area of the stock has previously been difficult to measure (Winters and Wheeler 1985), but VMS data can be used to approximate the total stock area with relative ease. Our estimates of gear efficiency varied depending on the selection criteria and study area. This uncertainty could be accommodated in a stock assessment in the same way that uncertainty in other parameters is commonly handled. The impact of various estimates of gear efficiency could be evaluated through a sensitivity analysis. Stock assessment forecasts could be run in a stochastic manner where gear efficiency estimates are chosen from the plausible range identified in this study.

We have calculated q for the entire fishery (Table 11.4) by coupling three plausible gear efficiency estimates with estimates of total stock area (calculated from one-minute scale annual effort data for the scallop fishery obtained from the aggregation analysis (i.e. areas accounted for the top 99% of total annual catch for each year). Our q estimates are effectively rescaled through the application of different E values, but the temporal trends in catchability are consistent between series (high in 2003 and low in 2001) due to the change in the size of total stock area.

Table 11.4 Approximate total stock area and catchability coefficients with different E estimates from fishing year 2001–2005.

Fishing year	Approximate stock area (nm^2)	q (E = 0.5)	q (E = 0.6)	q (E = 0.7)
2001	4710	0.66×10^{-5}	0.80×10^{-5}	0.93×10^{-5}
2002	3879	0.81×10^{-5}	0.97×10^{-5}	1.13×10^{-5}
2003	3018	1.04×10^{-5}	1.24×10^{-5}	1.45×10^{-5}
2004	3325	0.94×10^{-5}	1.13×10^{-5}	1.32×10^{-5}
2005	3179	0.98×10^{-5}	1.18×10^{-5}	1.38×10^{-5}

Constant or variable q?

High standard deviation observed in q estimates in the restrictive sensitivity run indicated that catchability is still highly variable among cells which have similar and high fishing effort. We propose two probable reasons for these inconsistencies:

- biological and environmental factors which affect the variability in the catchability of the stock; and/or
- errors in the data: difficulties in using TrackMapper to allocate catch to cells at a fine spatial resolution.

It is well recognised that there are many factors that affect fishing catchability especially for living marine resources with aggregated spatial distributions (Perez and Defeo 2003). Gonzalez-Yanez *et al.* (2006) summarised these factors into three groups:

1) biological factors such as size, sex, age, growth, lifecycle stage, migration, and fish behaviours (Joll and Penn 1990; Newby and Hansen 2000; Ralston and Tagami 1992)
2) environmental factors such as moon phases, temperature, seasonal changes, salinity (Hilborn and Walters 1992; Swain, Poirier *et al.* 2000; Wright, Caputi *et al.* 2006), and
3) the fishing factors such as fishing power (Bertrand, Diaz *et al.* 2004; Haddon 2001).

In this study, some biological factors such as migration and natural mortality were likely to be negligible due to the short study period and the intensity of the fishery, but other factors such as temperature, weather conditions, depth and sediment type were likely to differ from day to day and cell to cell, potentially causing variability in q (Weinberg, Rago *et al.* 2000).

We should be aware of the uncertainties in fine-scale fishing effort and CPUE data estimated from TrackMapper, which are identified as the spatial uncertainties of vessel positions and catch records. The uncertainty in the catch records, in particular, was expected to be high due to the difficulties in linking daily logbook catch records into hourly VMS trawl tracks. Peripheral areas have very low effort but can be assigned high catch records from tracks radiating out from high-density areas. This explains the high q values observed for peripheral cells in analyses done using the more relaxed selection criteria. The irony is that at smaller spatial scales, the more likely it is that assumptions of DeLury analysis will be met but the less likely it is that the catch assigned to the area will actually have been caught within that area. Unfortunately, there was not enough shot-by-shot logbook data to validate the results and we could not partition the inconsistencies in catchability into those resulting from error in the data and those resulting from other biological/environmental factors. Nevertheless, the aggregation and depletion analyses showed that the fishing effort is spatially aggregated and catchability is likely to be variable. The latest stock assessment of the Queensland scallop fishery (O'Neill, Courtney *et al.* 2005) corrected for biases caused by the spatially aggregated fishing effort and variable q in the catch rate data to some extent by standardising CPUE for fishing power and fishing area. However, the spatial scale of the fishing area remained relatively coarse (30 × 30-minute grid) and catchability was assumed to be constant despite the change in the spatial distribution of the fishing effort.

Although there are some issues in the spatial accuracy of the fishing effort and catches calculated from TrackMapper, we demonstrated that the VMS data can be used to improve future stock assessments. CPUE standardisation procedures (Chapter 10) will be improved through the identification of areas where fishing effort is concentrated and the application of appropriate spatial weighting to the data.

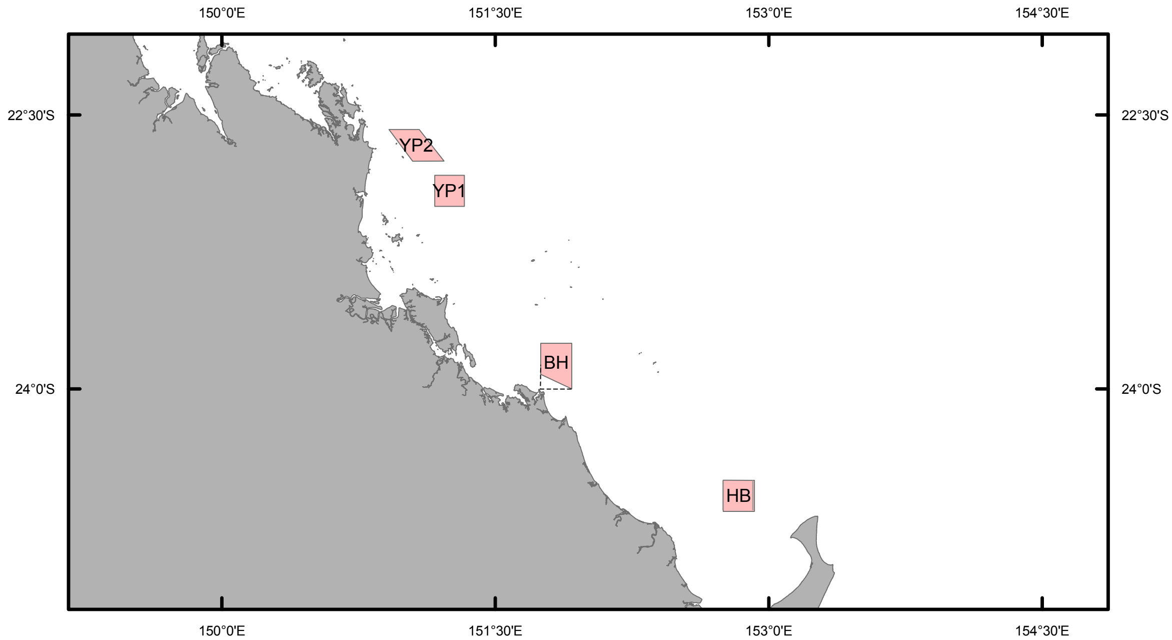


Figure 11.10. Original configuration of SRAs. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay. Note: The dashed line at the bottom of the Bustard Head closure indicates the original area.

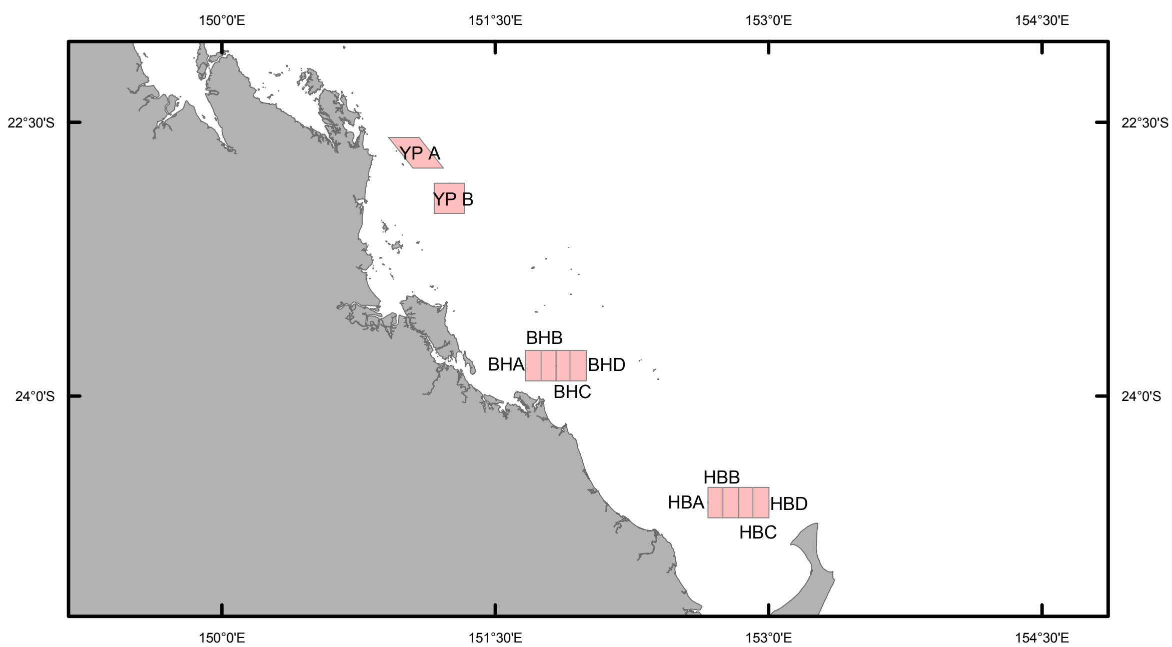


Figure 11.11. Configuration of SRAs following the implementation of the *Fisheries (East Coast Trawl) Management Plan 1999*. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay.

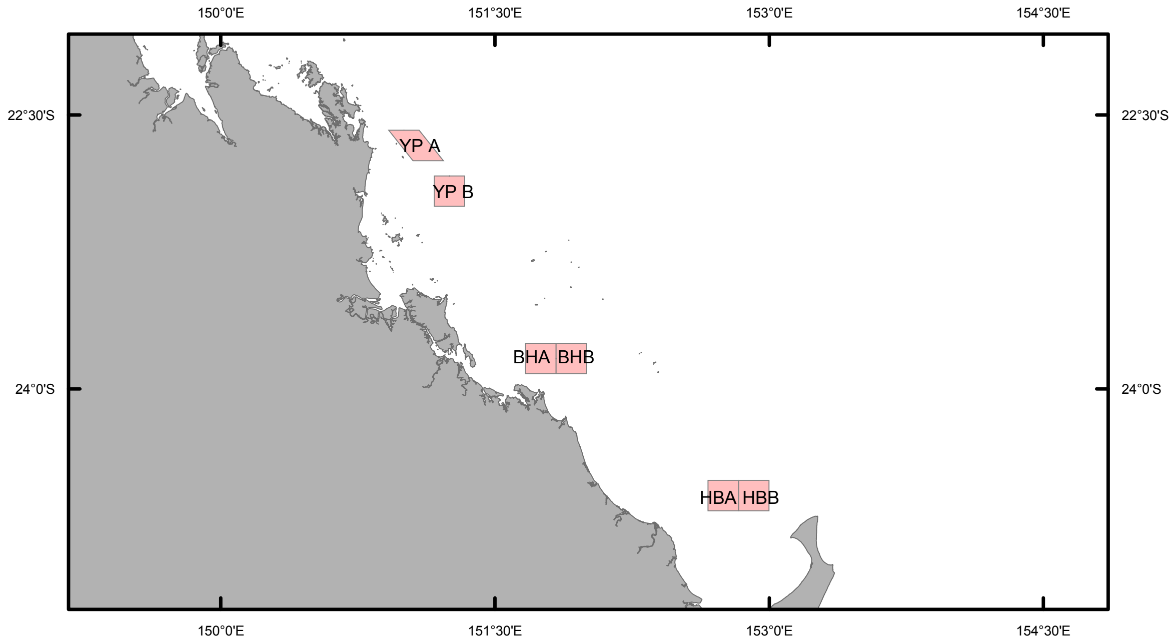


Figure 11.12. Current configuration of SRAs. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay.

Table 11.5. Periods of closure of original SRAs. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay.

SRA	Closed	Open
BH	11-Mar-97	1-Feb-01
HB	11-Mar-97	1-Feb-01
YP1	11-Mar-97	6-Mar-98
YP2	6-Mar-98	1-Feb-01

Table 11.6 Periods of closure of SRAs following the implementation of the *Fisheries (East Coast Trawl) Management Plan 1999*. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay.

SRA	Closed	Open	Closed
BHA	1-Feb-01	1-Jan-02	20-Sep-03
BHB	1-Feb-01	1-Jan-03	20-Sep-03
BHC	20-Sep-01	31-Oct-03	
BHD	20-Sep-02	31-Oct-03	
HBA	1-Feb-01	1-Jan-02	20-Sep-03
HBB	1-Feb-01	1-Jan-03	20-Sep-03
HBC	20-Sep-01	31-Oct-03	
HBD	20-Sep-02	31-Oct-03	
YPA	20-Sep-01	1-Jan-03	20-Sep-03
YPB	1-Feb-01	1-Jan-02	20-Sep-02

Table 11.7 Periods of closure of current SRAs. YP: Yeppoon, BH: Bustard Head, HB: Hervey Bay.

SRA	Closed	Open	Closed	Open	Closed
YPA	20-Sep-01	1-Jan-03	20-Sep-03	1-Jan-05	20-Sep-05
YPB	1-Feb-01	1-Jan-02	20-Sep-02	1-Jan-04	20-Sep-04
BHA			20-Sep-03	1-Jan-05	20-Sep-05
BHB			20-Sep-03	1-Jan-04	20-Sep-04
HBA			20-Sep-03	1-Jan-05	20-Sep-05
HBB			20-Sep-03	1-Jan-04	20-Sep-04

12 BENEFIT AND ADOPTION

As a concrete demonstration of the adoption by government of the output from this project, a fully functional version of our mapping software – TrackMapper – was operationalised within the secure area of the QDPI&F VMS unit in June 2005.

So far the project has made significant contributions to the management of the Queensland trawl fishery in three main areas. The first is a State Government response to the draft Representative Areas Program undertaken by the GBRMPA. The main output is a map of total effort at one-minute scale (approximately one square nautical mile) for the entire Queensland coast. This map ensured a level playing field for all stakeholders in determining the final boundaries of the review of Great Barrier Reef Marine Park Zoning Plan 2003.

The second was an internal Trawl Effort Review. It was a requirement under the legislated *Trawl Management Plan* to synthesise information from a number of research projects to review the allocation of effort within the East Coast Trawl Fishery. One of the objectives was to accurately define the extent of effort in respective fisheries at the one-minute scale. The resulting map accurately identified critical trawl grounds for all sectors.

The third was to provide a report to the Australian Government Department of the Environment and Heritage (DEH) detailing the percentage of trawl effort and amount of Gross Value of Production removed from the Great Barrier Reef Marine Park as a result of the introduction of the RAP as of 1 July 2004. This report was instrumental in developing the Structural Readjustment Package for the East Coast Otter Trawl Fishery (ECOTF) because of the RAP's introduction .

In addition, outputs from this project have been used in a number of applications by various agencies:

- For use in the Grey Nurse Regulatory Impact Statement to identify trawl effort in proposed grey nurse sanctuaries (QDPI&F);
- Identifying potential oyster mariculture sites within Moreton Bay (QDPI&F);
- Trawl-density maps to update survey design for *Seabed Biodiversity on the Continental Shelf of the Great Barrier Reef World Heritage Area* (CRC Reef, AIMS, QDPI&F, Queensland Museum and CSIRO; FRDC part-funded project**,** CRC REEF C1.1.2A);
- Identifying the spatial extent and separation of tiger and endeavour prawn stocks in the Torres Strait (QDPI&F, Fisheries Resources Research Fund and CRC Torres Strait);
- Identifying the distribution of sea snakes in the Queensland east coast (QDPI&F FRDC-funded project # 2005/053), and
- Various in-house maps for survey design in the Queensland ECOTF (QDPI&F).

As outlined in the project proposal, these examples demonstrated a ‘phased output’ of high-resolution maps of fishing effort and advice on spatial patterns in effort and catch to industry and management. This output was of critical benefit to the Queensland Otter Trawl Fishery in negotiations with GBRMPA and the Australian Government during the introduction of the RAP and in developing the Structural Readjustment Package. The readjustment package will provide an estimated $100 million to those affected in the fishing/seafood industry.

In recognition of the quality of these outputs, and their timeliness, the VMS project received the Queensland Seafood R&D Award from the Queensland Seafood Industry Association on 17 June 2005.

The benefit to other stakeholders is more difficult to quantify. The VMS project outputs have been requested and used by both Australian and State Government agencies for fisheries management and conservation, and they have been used by other research projects in the design of surveys or analytical protocols. In this regard, one of the major beneficiaries of the VMS project output has been the multi-institutional *Seabed Biodiversity on the Continental Shelf of the Great Barrier Reef World Heritage Area* project. A component of the $7 million *Seabed biodiversity* project, aside from providing seminal baseline information for the GBR, is to assess the impact of trawling. Fine spatial resolution data on trawl effort supplied by the VMS project match the survey resolution. This will therefore maximise the accuracy and benefit of the ecological impact assessment.

A further benefit, although more intangible, has been the dissemination of the concept of using VMS data in resource assessment, not just for compliance. The project team has presented the methodology and results of the project at state, national and international venues. Requests for information on the project have come from Europe via the International Oceans Institute, the USA via National Oceanic and Atmospheric Administration, and from Asia via the International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences, held in Shanghai, China.

13 FURTHER DEVELOPMENT

This project experienced critical difficulties due to data deficiencies related to the assignment of:

- activities to vessel tracks, and
- catch to trawl tracks.

Improved definition of vessel activity is the most tractable problem. Whilst our Hidden Markov Modelling approach provided a technically elegant solution, it achieved only marginal increases in accuracy over more parsimonious filter methods of identifying trawl tracks. The need for these estimation procedures could be obviated altogether by directly recording data that accurately characterises vessel activity. The deployment of research logbooks and observers may allow accurate recording for sampled components of the fleet but may not achieve broad coverage. A more complete validation of vessel activity could be achieved through technological solutions. In some Australian fisheries continuously recorded camera footage is used successfully to monitor deck activity (Bruce Wallner (AFMA), pers. comm.). Another possibility that may achieve even wider application could be the use of tension meters to log trawl warp tension – a key indicator of trawling activity.

Improving the spatio-temporal resolution at which catch is recorded is more problematic. The initial design of the project was predicated on the availability of data from ECERS. Whilst ECERS would have provided timely data in an efficient manner it would still have suffered from the same problems when assigning catch recorded in logbooks to VMS-derived trawl tracks. ECERS records would probably still have been reported each daily, rather than shot by shot. Even when shot-by-shot logbook catch data were available to this project we still encountered difficulties in assigning catch to tracks because of the difference in spatial scales at which the catch and track data were available. Commercial trawl tows frequently extend over several hours. This makes it difficult to determine at which points along the trawl track particular volumes of catch were actually caught.

Shortened tow durations, coupled with shot-by-shot catch recording, would immediately improve resolution. However, these practices fundamentally interfere with normal fishing operations and cannot be practically implemented across a commercial fishing fleet. A more fruitful solution may be to collect high-resolution catch and effort data specifically to empirically validate the precision provided by low resolution commercial data. Several technological solutions may be particularly useful:

- Sophisticated catch monitoring systems (e.g. the Scanmar Trawleye catch monitoring system) can log real-time sonar images showing fish as they enter the trawl. Such images could be used to partition the catch from long tows into shorter time intervals.
- The use of a modified trawl fitted with a ‘multisampler’ cod end could further validate this approach. These devices allow multiple cod ends to

be opened and closed during a tow so that several separate samples of catch can be obtained from a single trawl deployment (Engås, Skeide *et al.* 1997).

When these critical data deficiencies are overcome the potential improvements to stock assessment resulting from the use of VMS data may be fully realised. We then expect that CPUE standardisation models will be greatly improved through the inclusion of VMS-derived parameters, and a longer time series of VMS data.

The availability of VMS information may also allow the development of alternative measures of the performance of fisheries. For example, the expansion or contraction of the area fished or changes in the concentration of effort are readily measured using VMS information and may be related to changes in the spatial coverage and productivity of the stock. The development of such VMS-related performance measures (and the choice of appropriate thresholds) requires further study of the relationship between these measures and stock productivity.

14 PLANNED OUTCOMES

This project sought to empower the Queensland trawl industry and fishery managers to meet present and future challenges by providing three main enhancements:

1) Better information about the status and sustainability of the resource

The fine-spatial resolution maps of catch and effort (status of the resource) we produced progressively over the life of the project have been used in a number of applications for various agencies. The use of these outputs from the VMS project represents a significant outcome for the project:

- For use in the Grey Nurse Regulatory Impact Statement to identify trawl effort in proposed grey nurse sanctuaries (QDPI&F);
- Identifying potential oyster mariculture sites within Moreton Bay (QDPI&F);
- One-minute effort maps for three fisheries (saucer scallop, EKP and North Queensland tiger prawn) for estimating the sustainability of the resource in the QDPI&F Trawl Effort Review (QDPI&F);
- Input via VMS-derived maps of the spatial distribution of trawl effort for the GBRMPA draft Representative Areas Program, to identify potential impact on trawl effort within Great Barrier Reef Marine Park (QDPI&F and Queensland Seafood Industry Association (QSIA));
- Maps detailing the amount of GVP lost as a result of the introduction of the RAP. This work was done for DEH to aid in developing a structural readjustment package to compensate fishers. It also involved discussions with the Queensland Seafood Industry Association in order to validate our methods for producing the estimates of lost GVP. (QDPI&F, QSIA, GBRMPA); and
- Information to update survey design for ‘Impacts of trawling on benthos’ study (CSIRO).

Furthermore, in 2004 and 2005 the project team had discussions with officers from NOAA Fisheries in the USA regarding a possible collaboration between countries to help them develop trawl signatures and trawl tracks using their VMS in their southern Rock Shrimp (*Sicyonia brevirostris*) fishery.

2) Reliable information on the distribution of trawled and untrawled areas

Seminars we have given to provide information on the spatial distribution of trawl catch and effort, and to promote and discuss the use of VMS data in natural resource assessment and management, include:

- Talks at Southern Fisheries Centre seminar series. Norm Good’s talk was titled ‘Mapping fishery resource intensity using Maximum Entropy modelling – or cleaning up dirty pictures‘, and David Peel’s talk was titled, ‘Developing trawl track and trawl signatures using VMS and GPS data’

- Norm Good also gave a talk summarising the latest results to the Fisheries Business Group of QDPI&F and to the Scientific Advisory Committee of Queensland TrawlMAC.
- David Peel gave an overview of the project to a delegation from the Zhejiang Provincial Oceanic & Fisheries Bureau, China.

An article was published in the February 2003 issue of the *Queensland Fisherman*, outlining the main objectives of the project and calling for help in providing GPS data for defining and validating trawl signatures and tracks. This article resulted in a favourable response in the next issue of *Queensland Fisherman* from the Trawl Chair of the Queensland Seafood Industry Association, Robin Hansen.

ABC News interviewed Norm Good and David Peel in June 2004. The story, about two minutes long, covered the effort mapping and stock assessment aspects of the project. A copy of the news vision was sent to Dr Patrick Hone (FRDC) on CD.

A half-page article published in the July 2004 issue of the *Professional Fisherman* gave a summary of the projects objectives and included an example map of trawl effort at the one-minute scale.

On 17 June 2005, the project received the Queensland Seafood R&D Award. The award was in recognition of the project's efforts in supplying the Queensland Seafood Industry Association with high-resolution fishing effort maps (based on our VMS data analysis) during the negotiations over the RAP.

On 20 August 2005, Norm Good presented a talk entitled 'An index of abundance for prawn stocks in Queensland using Maximum Entropy methods and VMS data' at the 3rd International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences, held in Shanghai, China. The presentation was awarded the 'best oral presentation' and a paper was published in the proceedings of the symposium.

On 2 November 2005, Neil Gribble presented a talk entitled 'Innovative fisheries resource assessment and fishing effort mapping using satellite-based VMS' at the IMarEST/IOI 'Peace in the Ocean' Conference, Townsville. The talk was published as an article in the IOI (International Oceans Institute) annual report on the Status of the Oceans.

An article was written for the FRDC *R&D News*, October 2005 outlining the current research outputs of the project and noting the Queensland Seafood R&D Award.

In 2005 a comprehensive briefing memo was sent to Bev Tyrer, QDPI&F Trawl Manager, summarising the main features of the software and how to obtain map outputs, including the distribution of trawled and untrawled areas, for circulation to all internal QDPI&F stakeholders.

3) Tools to help it make informed strategic decisions

Adoption (operationalisation) of the VMS project research outputs as a standard function by the QDPI&F VMS unit (see Appendix D for the software use guide).

30 June 2005: An initial meeting was held with the Manager (Bev Tyrer) and Senior Systems Analyst (Cameron Baker) of the QDPI&F VMS unit. Due to the confidential nature of the VMS data it was suggested that a fully functional version of the projects mapping software (termed TrackMapper) be held within the secure area of the unit. Requests for map products would be made through and approved by the Manager.

29 November 2005: The mapping software requires the integration of two databases. The daily logbook fishing records are held in the CFISH database administered by the DPI&F Assessment & Monitoring unit (Logbook unit). The VMS position information is held in the Traffic database administered by the QDPI&F VMS unit.

Note: A data warehousing strategy is being developed by QDPI&F to streamline the process of obtaining fishery-related information. A meeting was held with the Manager of the Logbook unit (Jim Higgs) and the VMS unit chief programmer Cameron Baker to define data requirements for updating the TrackMapper database. These data requirements were then incorporated into the user specifications of the data warehousing strategy. Once the strategy is in place a program called SetUpTrackMapper (written by our project team) may be run at any time to update the TrackMapper database. Until such time as the 'Warehousing strategy' comes into effect any database updates will be conducted by the project team as needed.

7 December 2005: A half-day workshop was held with relevant stakeholders from the Trawl, Logbook and VMS units to demonstrate the capabilities of the TrackMapper software and to canvas user requirements to further enhance its functionality. Outcomes from the workshop regarding functionality have been incorporated into the software, namely calculating the number of individual vessels trawling within a predefined grid. This will enable the generation of maps that can be made suitable for public viewing by restricting output to only those areas where at least five vessels are operating. Additionally, fisheries definitions for pipefish and bug species were developed for inclusion in to the software. As these species are byproduct the trawl signature definitions for EKP and saucer scallops respectively were used to filter records.

13 January 2006 – 10 February 2006: Installing and testing TrackMapper software and Oracle database on a standalone PC in the VMS unit.

10 February 2006: Software demonstrated to the QDPI&F VMS unit staff, Trawl Manager (Mark Lightowler), Senior Fisheries Economist (Lew Williams) and Fisheries and Boating Patrol Intelligence Analyst (Erica Ross). As well as the demonstration, there was considerable discussion regarding additional

capabilities to be incorporated into the software and/or research avenues. In particular Lew Williams was interested in calculating trip length and time away from port for economic analysis, and Erica Ross was interested in fine-scale (localised fishing aggregations) trawl pattern distributions.

A copy of the installation guide and user guide for TrackMapper program, used as standard operating procedure by QDPI&F, was forwarded to FRDC.

The project has also provided other resource management agencies such as DEH, with information about the utility of this novel approach to fishery assessment, and tools that may be adapted to their own resource assessments.

15 CONCLUSIONS

Objective 1: Review applications and potential of VMS mapping and OceanFARM software, and related approaches.

- At the beginning of the project there were relatively few published studies relevant to Queensland fisheries.
- In the 2005–2006 period however there have been a number of published final reports and journal articles that are relevant to this project. These have identified the application of the VMS data to the mapping of fishing effort (e.g. the prawn trawl fishery of the Gulf of Carpentaria) and the application of Maximum Entropy methods to fine-spatial resolution survey data.

Objective 2: Develop trawl track and trawl signature definitions for each fishery sector to use with TerraVision software.

- VMS TrackMapper software developed by the project is now standard operating procedure for the QDPI&F VMS unit.
- Decision rules using a statistical mixture model to identify speed cut-off points for trawl signatures in each trawl fishery can accurately identify the majority of trawl behaviour.
- The Hidden Markov Model (HHM) method is statistically more rigorous, more objective, and better describes the errors in trawl signature identification. However, whilst a marginal improvement in accuracy was achieved, the method requires further validation.
- In the interest of parsimony, the simpler filter method using the mixture model approach has been used in the operational version of TrackMapper software (adopted by the QDPI&F VMS unit).

Objective 3: Map the spatial and temporal intensity of fishing effort in each trawl sector, and estimate the distribution and extent of trawled and untrawled areas.

- Maps were produced of spatial intensity of fishing effort for each major trawl fishery along the Queensland east coast at a spatial resolution of one nm and with a minimum temporal resolution of one day.
- From these maps it was possible to accurately identify the distribution and extent of trawled and untrawled areas (these maps were requested by QSIA and GBRMPA in negotiations over the positioning of RAP and have been used subsequently in compensation negotiations with the fishing industry).

Objective 4: Map resource density indices for each fishing sector.

- Maximum Entropy analysis (MaxEnt) was applied to VMS effort data combined with CFISH logbook catch data to produce exploratory density indices for EKP in Queensland.
- While showing great potential, the method highlighted that effort and catch data must be at similar high-spatial resolution to get an accurate density estimate.
- There was also a problem of temporal blurring if the effort data were collected over a short time period (duration of trawl) but the catch records were integrated over a longer period (a night's fishing, effectively 24 hours).
- The MaxEnt algorithm appears to need data on tracks with zero catches to adequately describe the underlying density indices. There is a potential mismatch because the current QDPI&F logbook database records catches but does not record the proposed target species hence cannot record zero catches.

Objective 5 (as amended): Assess the trawl fishery management plan Review Events and other reference points given the improved spatial definition of the data. Develop improved data inputs for stock assessments, using the Eastern King Prawn stock as a case study.

- Maps generated by the project and by the VMS unit have been used to help to describe areas fished, and areas now protected relative to areas previously fished, in negotiations for compensation of trawl fishers in industry restructuring.
- Fine detail mapping of VMS information was incorporated into the current standardisation model for the EKP fishery and used to evaluate CPUE-related reference points.
- Fine detail mapping of VMS information was incorporated into Stock Assessment of the Torres Strait Tiger Prawn Fishery (*Penaeus esculentus*). This was a major review event prompted by legislative change to the management plan.
- Using VMS for depletion estimates showed promise but also showed areas of difficulty.
- As with the Maximum Entropy analysis, the depletion study highlighted that the scales and resolution of effort and catch data must be similar to get an accurate depletion estimate.
- For the all three fisheries explored, fishing effort was found to be spatially aggregated:
    - Analysis of Tiger/endeavour prawn fishery VMS data suggested that targeting occurred in areas of high CPUE; interpreted as fishers targeting aggregations of prawns
    - Analysis of the Scallop and EKP fisheries, in contrast, showed that although effort was aggregated it was not related to areas of high

CPUE; interpreted as fishers following spatial patterns determined by external processes which might include management closures, fuel prices, and cost-benefit business decisions.

- Future calculation of CPUE indices should involve the application of appropriate spatial weightings to correct for the concentration of fishing effort within the fishery area, and for the targeting of areas with higher catch rates.
- It was apparent from both the depletion and aggregation analyses that the limiting factor was the spatial and temporal resolution of the catch data. Greater resolution in the logbook records of the catch would be improved by shot-by-shot information (preferably via electronic logbooks as per the original proposal).

16 REFERENCES

Afonso-Dias M (2000) Use of satellite GPS data to map effort and landings of the Portuguese crustacean fleet. *ICES CM 2000/G:06*.

Afonso-Dias M, Simões J, Pinto C (2002) A dedicated GIS to estimate and map fishing effort and landings per unit effort for the Portuguese crustacean trawl fleet *(Unpublished)*.

Arreguin-Sanchez F (1996) Catchability: A key parameter for fish stock assessment. *Reviews in Fish Biology and Fisheries* **6**, 221-242.

Baum LE, Peterie T, Souled G, Weiss N (1970) A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics* **41**, 164-171.

Begg GA, Marteinsdottir G (2003) Spatial partitioning of relative fishing mortality and spawning stock biomass of Icelandic cod. *Fisheries Research* **59**, 343-362.

Bene C, Tewfik A (2001) Fishing effort allocation and fishermen's decision making process in a multi-species small-scale fishery: Analysis of the conch and lobster fishery in Turks and Caicos Islands. *Human Ecology* **29**, 157-186.

Bertrand S, Burgos JM, Gerlotto F, Atiquipa J (2005) Lévy trajectories of Peruvian purse-seiners as an indicator of the spatial distribution of anchovy (*Engraulis ringens*). *ICES Journal of Marine Science* **62**, 477-482.

Bertrand S, Diaz E, Niquen M (2004) Interactions between fish and fisher's spatial distribution and behaviour: an empirical study of the anchovy (*Engraulis ringens*) fishery of Peru. *ICES Journal of Marine Science* **61**, 1127-1136.

Beukers-Stewart BD, Jenkins SR, Brand AR (2001) The efficiency and selectivity of spring-toothed scallop dredges: A comparison of direct and indirect methods of assessment. *Journal of Shellfish Research* **20**, 121-126.

Bishop J, Venables WN, Wang Y-G (2004) Analysing commercial catch and effort data from a Penaeid trawl fishery. A comparison of linear models, mixed models, and generalised estimating equations approaches. *Fisheries Research* **70**, 179-193.

Bontekoe TR, Koper E, Kester DJM (1994) Pyramid maximum entropy images of IRAS survey data. *Astronomy and Astrophysics* **284**, 1037-1053.

Booth AJ (2000) Incorporating the spatial component of fisheries data into stock assessment models. *ICES Journal of Marine Science* **57**, 858-865.

Brierley AS, Gull SF, Wafy MH (2003) A Bayesian maximum entropy reconstruction of stock distribution and inference of stock density from line-transect acoustic-survey data. *ICES Journal of Marine Science* **60**, 446-452.

Burg JP (1975) Maximum entropy spectral analysis. Ph.D. thesis. Stanford University, USA. (Unpublished).

Burridge CY, Pitcher CR, Hill BJ, Wassenberg TJ, Poiner IR (2006) A comparison of demersal communities in an area closed to trawling with those in adjacent areas open to trawling: A study in the Great Barrier Reef Marine Park, Australia. *Fisheries Research* **79**, 64-74.

Burridge CY, Pitcher CR, Wassenberg TJ, Poiner I, Hill BJ (2003) Measurement of the rate of depletion of benthic fauna by prawn(shrimp) otter trawls: an experiment in the Great Barrier Reef, Australia. *Fisheries Research* **60**, 237-253.

Campbell RA (2004) CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fisheries Research* **70**, 209-227.

Campbell RA, Tuck G, Tsuji S, Nishida T (1996) Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. In 'Working Paper SBFWS/96/16 Presented at the Second CCSBT Scientific Meeting'. Hobart, Australia

Courtney AJ, Masel JM, Die DJ (1995) Temporal and spatial patterns in recruitment of three penaeid prawns in Moreton Bay, Queensland, Australia. *Estuarine, Coastal and Shelf Science* **41**, 377-392.

Cowx IG (1983) Review of the methods for estimating fish population size from survey removal data. *Fisheries Management* **14**, 67-82.

DeLury DB (1947) On the estimation of biological populations. *Biometrics* **3**, 145-167.

Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1-38.

Deng R, Dichmont C, Milton D, Haywood M, Vance D, Hall N, Die D (2005) Can vessel monitoring system data also be used to study trawling intensity and population depletion? The example of Australia's northern prawn fishery. *Canadian Journal of Fisheries & Aquatic Sciences* **62**, 611-622.

Dichmont CM, Die D, Punt AE, Venables W, Bishop J, Deng A, Dell Q (2001) 'Risk analysis and sustainability indicators for prawn stocks in the Northern Prawn Fishery. Final Report on FRDC Project No. 1998/109.' CSIRO Marine Research, Cleveland.

Dichmont CM, Haddon M, Yeomans K, Kelly K (1999) 'Proceedings of the South-East Queensland Stock Assessment Review Workshop.' Department of Primary Industries, QC99003, Brisbane, Queensland, Australia.

Dorn MW (2001) Fishing behavior of factory trawlers: a hierarchical model of information processing and decision-making. *ICES Journal of Marine Science* **58**, 238-252.

Engås A, Skeide R, West CW (1997) The "Multisampler": a system for remotely opening and closing multiple codends on a sampling trawl. *Fisheries Research* **29**, 295-298.

Fletcher WJ (1992) Use of a spatial model to provide initial estimates of stock size for a purse seine fishery on pilchards (*Sardinops sagax neopilchardus*) in Western Australia. *Fisheries Research (Amsterdam)* **14(1)**, 41-57.

Francis RICC, Hurst RJ, Renwick JA (2003) Quantifying annual variation in catchability for commercial and research fishing. *Fishery Bulletin* **101**, 293-304.

Gedamke T, DuPaul WD, Hoenig JM (2004) A spatially explicit open-ocean DeLury analysis to estimate gear efficiency in the Dredge Fishery for sea scallop *Placopecten magellanicus*. *North American Journal of Fisheries Management* **24**, 335-351.

GENSTAT 5 Committee (1993) 'Genstat 5 release 2 reference manual.' (ClarendonPress: Oxford)

GENSTAT (2005) GENSTAT 8th Edition. In. (Lawes Agricultural Trust)

Gonzalez-Yanez AA, Millan RP, de Leon ME, Cruz-Font L, Wolff M (2006) Modified DeLury depletion model applied to spiny lobster, *Panulirus argus* (Latreille, 1804) stock, in the southwest of the Cuban Shelf. *Fisheries Research* **79**, 155-161.

Good N, Peel D (2005) An index of abundance for prawn stocks in Queensland using Maximum Entropy methods and VMS data. *Proceedings of the Third International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences*, (2005) In Press.

Gribble NA, Peel D, Good N (2006) Innovative fisheries resource assessment and fishing effort mapping using satellite based VMS. Proceedings from IMarEST/IOI. IOI(International Oceans Institute) annual report on the Status of the Oceans. In 'Proceedings of IMarEST/IOI Pacem in Maribus XXXI conference'. Townsville, pp506. (Eds GR South and C Boese)

Gull SF, Daniell GJ (1978) Image reconstruction from incomplete and noisy data. *Nature* **272**, 686-690.

Gull SF, Skilling J (1999) Quantified maximum entropy. MemSys5 User's Manual, V.1.2. In. (Maximum Entropy Data Consultant Ltd: Suffolk, UK)

Gulland JA (1983) 'Fish Stock Assessment: A manual for basic method.' (John Wiley and Sons: New York)

Haddon M (2001) 'Modelling and Quantitative Methods in Fisheries.' (Chapman and Hall: New York)

Haddon M, Harrington JJ, Semmens JM (2006) 'Juvenile scallop discard rates and bed dynamics: testing the management rules for scallops in Bass Strait. Final Report on FRDC Project No. 2003/017.' Fisheries Research and Development Corporation and Tasmanian Aquaculture and Fisheries Institute, Hobart.

Hall NJ, Vance DJ, Haywood MDE, Die DJ (1999) Monitoring the small-scale distribution of fishing intensity in a tropical prawn trawl fishery. *(Unpublished)*.

Hampton J, Fournier DA (2001) A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine & Freshwater Research* **52**, 937-963.

Harrington J, Semmens J, Haddon M (2006) Comparison of dredged benthic fauna from areas of different fishing intensity - examining the effects of fishing disturbance. *Journal of Shellfish Research* **25**, 285.

Haywood M, Hill B*, et al.* (2005) Quantifying the effects of trawling on seabed fauna in the Northern Prawn Fishery.

Heywood BG, Brierley AS, Gull SF (2006) A quantified Bayesian Maximum Entropy estimate of Antarctic krill abundance across the Scotia Sea and in small-scale management units from the CCAMLR-2000 survey. *Journal of the Scientific Committee and Commission for the Conservation on Antarctic Living Resources* **13**, 97-116.

Hilborn R, Walters CJ (1992) 'Quantitative Fisheries Stock Assessment: Choice Dynamics and Uncertainty.' (Chapman and Hall: New York)

Hilborne R, Ledbetter M (1979) Analysis of the British Columbia Salmon Purse-sein Fleet; Dynamics of Movement. *Canadian Fisheries Resources Board* **36**, 384-391.

Jaynes ET (1957) Information and statistical mechanics. *Physical Review* **106**, 620-630.

Jaynes ET (1980) The minimum entropy production principle. In 'Annual review of physical chemistry'. (Ed. R Rosenkranz) pp. 579-601 Reprinted in Jaynes ET,

(1983) Papers on probability, statistics and statistical physics. (D. Reidel Publishing Co.: Dordrecht-Holland)

Joll LM, Penn JW (1990) The application of high-resolution navigation systems to Leslie-DeLury depletion experiments for the measurement of trawl efficiency under open-sea conditions. *Fisheries Research* **9**, 41-55.

Kangas M, Sporer E, Harvey D (2005) Broome Prawn Managed Fishery Status Report. In 'State of the fisheries Report 2003/04'. (Eds JW Penn, WJ Fletcher and F Head) pp. 118-120. (Department of Fisheries: Western Australia)

Koslow JA, Gowlett-Holmes K, Lowry JK, O'Hara T, Poore GCB, Williams A (2001) Seamount benthic macrofauna off southern Tasmania: community structure and impacts of trawling. *Marine Ecology-Progress Series* **213**, 111-125.

Larcombe WP, McLoughlin K, Tilzey RDJ (2001) Trawl operations in the South East Fishery, Australia: spatial distribution and intensity. *Marine & Freshwater Research* **52**, 419-30.

Lizamore S (1995) Topics in Maximum Entropy Applications. Master's Thesis. ISOR, Victoria University of Wellington, New Zealand.

Llacer L, Núñez J (1990) Iterative maximum likelihood estimator and Bayesian algorithms for image reconstruction in astronomy. In 'The restoration of HST images and spectra'. (Eds RL White and RJ Allen) pp. 62-70. (Space Telescope Science Institute: Baltimore)

MacDonald IL, Zucchini W (1997) 'Hidden Markov and other models for discrete-valued time series.' (Chapman & Hall: London)

Maravelias CD, Reid DG, Simmonds EJ, Haralabous J (1996) Spatial analysis and mapping of acoustic survey data in the presence of high local variability: geostatistical application of North Sea herring (*Clupea harengus*). *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 1497-1505.

Marrs SJ, Tuck ID, Atkinson RJA, Stevenson TDI, Hall C (2002) Position data loggers and logbooks as tools in fisheries research: results of a pilot study and some recommendations. *Fisheries Research* **58**, 109-117.

Maunder MN, Punt AE (2004) Standardizing catch and effort data: a review of recent approaches. *Fisheries Research* **70**, 141-159.

Maury O, Gascuel D (2001) 'Local overfishing' and fishing tactics: theoretical considerations and applied consequences in stock assessment studied with a numerical simulator of fisheries. *Aquatic Living Resources* **14**, 203-210.

Maury O, Gascuel D, Fonteneau A (2001) Spatial modelling of Atlantic Yellowfin Tuna population dynamics: Application of a habitat-based advection-diffusion-reaction model to the study of local overfishing. In 'Spatial processes and management of marine populations.' (Eds GH Kruse, N Bez, A Boothet al). (University of Alaska Sea Grant: Fairbanks)

McLachlan GJ, Peel D (2000) 'Finite mixture models.' (John Wiley and Sons: New York)

Mejias. A (1999) Vessel monitoring system sensor applications in the Gulf of Mexico shrimp fishery. In 'Proceedings of the International Conference on Integrated Fisheries Monitoring.' Sydney, Australia. (Ed. CP Nolan). (FAO 1999)

Murawski SA, Wingley SE, Fogarty MJ, Rago PJ, Mountain DG (2005) Effort distribution and catch patterns adjacent to temperate MPAs. *ICES Journal of Marine Science* **62**, 1150-1167.

Murray AWACogarssaoAkadIJoMS- (1996) Comparison of geostatistical and random sample survey analyses of Antarctic krill acoustic data. *ICES Journal of Marine Science* **53**, 415-421.

Newby JR, Hansen MJ (2000) Catchability of walleyes to angling in Escanaba Lake, Wisconsin, 1980 -1995. *North American Journal of Fisheries Management* **20**, 873-881.

Nishida T, Booth AJ (2001) Recent approaches using GIS in the spatial analysis of fish populations. In 'Spatial processes and management of marine populations.' Fairbanks. (Eds GH Kruse, N Bez, A Boothet al) pp. 19-36. (University of Alaska Sea Grant)

O'Neill MF, Courtney AJ, Good NM, Turnbull CT, Yeomans KM, Staunton Smith J, Shootingstar C (2005) 'Reference point management and the role of catch-per-unit effort in prawn and scallop fisheries. Final Report on FRDC Project No. 1999/120.' Department of Primary Industries and Fisheries, QO05001, Brisbane, Queensland, Australia.

O'Neill MF, Courtney AJ, Turnbull CT, Good NM, Yeomans KM, Smith JS, Shootingstar C (2003) Comparison of relative fishing power between different sectors of the Queensland trawl fishery, Australia. *Fisheries Research* **65**, 309-321.

O'Neill MF, Leigh GM (2006) 'Fishing power and catch rates in the Queensland east coast trawl fishery.' Department of Primary Industries and Fisheries, QI06051, Brisbane, Queensland, Australia.

Paloheimo JE, Dickie LM (1964) Abundance and fishing success. *Rapports et Proces-Verbaux des Reunions, Conseil International Pour l'Exploration de la Mer* **155**, 152-163.

Perez EPE, Defeo O (2003) Time-space variation in the catchability coefficient as a function of catch per unit of effort in *Heterocarpus reedi* (Decapoda, Pandalidae) in north-central Chile. *Interciencia* **28**, 178-182.

Pet-Soede C, Van Densen WLT, Hiddink JG, Kuyl S, Machiels MAM (2001) Can fishermen allocate their fishing effort in space and time on the basis of their catch rates? An example from Spermonde Archipelago, SW Sulawesi, Indonesia. *Fisheries Management and Ecology* **8**, 15-36.

QECTMP (2004) 'Fisheries (East Coast Trawl) Management Plan 1999; Fisheries Act 1994.'

Rabiner LR (1989) A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of IEEE, 77, 257-286.

Ralston S, Tagami DT (1992) An assessment of the exploitable biomass of *Heterocarpus laevigatus* in the main Hawaiian Islands. Part 1: Trapping surveys, depletion experiment, and length structure. *Fishery Bulletin* **90**, 494-504.

Ricker WE (1975) Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada* **191**, 382pp.

Rijnsdorp AD, Buys AM, Storbeck F, Visser EG (1998) Micro-scale distribution of beam trawl effort in the southern North Sea between 1993 and 1996 in relation to the trawling frequency of the sea bed and the impact on benthic organisms. *ICES Journal of Marine Science* **55**, 403-419.

Rijnsdorp AD, Dol W, Hoyer M, Pastoors MA (2000) Effects of fishing power and competitive interactions among vessels on the effort allocation on the trip level of the Dutch beam trawl fleet. *Ices (International Council for the Exploration of the Sea) Journal of Marine Science* **57**, 927-937.

Rijnsdorp AD, van Mourik Broeman PL, Visser EG (2000) Competitive interactions among beam trawlers exploiting local patches of flatfish in the North Sea. *Ices (International Council for the Exploration of the Sea) Journal of Marine Science* **57**, 894-902.

Rivoirard J, Simmonds J, Foote KG, Fernandes P, Bez N (2000) 'Geostatistics for estimating fish abundance.' (Blackwell Science: Oxford)

Sanders MJ (1988) Mean population number and the De Lury and Leslie methods. *Fisheries Research* **6**, 153-165.

Seber GAF (1982) 'The Estimation of Animal Abundance.' (Charles Griffin and Company LTD: London)

SHONO H (2005) Is model selection using Akaike's information criterion appropriate for catch per unit effort standardization in large samples? *Fisheries Science* **71**, 978-986.

Smith A (2001) Vessel Monitoring Systems, UN Atlas of the Oceans. In.

Swain DP, Poirier GA, Sinclair AF (2000) Effect of water temperature on catchability of Atlantic cod (*Gadus morhua*) to the bottom-trawl survey in the southern Gulf of St Lawrence. *ICES Journal of Marine Science* **57**, 56-68.

Swain DP, Sinclair AF (1994) Fish distribution and catchability: what is the appropriate measure of distribution? *Canadian Journal of Fisheries & Aquatic Sciences* **51**, 1046-1054.

Trimmer M, Petersen J, Sivyer DB, Mills C, Young E, Parker ER (2005) Impact of long-term benthic trawl disturbance on sediment sorting and biogeochemistry in the southern North Sea. *Marine Ecology Progress Series* **298**, 79-94.

Vignaux M, Vignaux GA, Lizamore S, Gresham D (1998) Fine-scale mapping of fish distribution from commercial catch and effort data using maximum entropy tomography. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1220-1227.

Weinberg J, Rago P, Keith C, Hendrickson L, Murawski S, Powell E, Mann R, Weidman C (2000) Stock assessment of surfclams along the east coast of the United States: the importance of estimating dredge efficiency. *Journal of Shellfish Research* **19**, 627.

Williams LE (2002) 'Queensland's Fisheries Resources: current condition and recent tends 1988-2000.' Department of Primary Industries, Queensland Department of Primary Industries, Information Series QI02012, Brisbane.

Winters GH, Wheeler JP (1985) Interaction between stock area, stock abundance, and catchability coefficient. *Canadian Journal of Fisheries & Aquatic Sciences* **42**, 989-998.

Wright I, Caputi N, Penn JW (2006) Depletion-based population estimates for western rock lobster (*Panulirus cygnus*) fishery in Western Australia. *New Zealand Journal of Marine and Freshwater Research* **40**, 107-122.

17 ACKNOWLEDGMENT

We would like to thank Bev Tyrer (VMS manager) and VMS unit staff for providing the VMS data. Thanks also go to the steering committee members:
Brigid Kerrigan (QDPI&F Trawl Fishery Manager), Geoff Tilton (Commercial fisher), Cameron Baker (QDPI&F Senior Systems Analyst VMS), Yimin Ye (CSIRO, Stock Assessment Scientist) for helpful feedback during the development phase of the project. We thank commercial fishers who volunteered to provide personal GPS track information to assist the validation process. Many thanks go to Dr. Todd Gedamke (NOAA Florida, Research Fisheries Biologist) for providing practical and theoretical advice on Depletion Analysis. We thank TerraVision staff for many helpful discussions particularly in the early phases of the project.

The project was funded by the Australian Fisheries Research and Development Corporation (FRDC). Thanks also go to Tony Courtney for his assistance in liaison in TrawlMac, Toby Patterson for the seminar that inspired the use of HMM in this work, and Mark Bravington for his insightful comments and suggestions.

18 APPENDICES

18.1 Appendix A: Project Staff

Staff engaged on the project were:

Simon Hoyle (Project Principal Investigator, 2002)
Neil Gribble (Project Principal Investigator, 2003–06)
Norm Good
David Peel
Mai Tanimoto
Rick Officer

18.2 Appendix B: Intellectual Property

It is not the intention of this project to commercialise any of the components of the research.

18.3 Appendix C: Confidentiality Deed

CONFIDENTIALITY DEED

Confidential Information of   of   (the 'Trawl Fisherman') may become known to the State of Queensland through the Department of Primary Industries ('DPI') by the Trawl Fisherman providing GPS data consisting mainly of location, time and depth information to DPI and its officers for use in a Fisheries Research and Development Corporation funded project entitled *'Innovative Stock Assessment and Effort Mapping Using VMS and Electronic Logbooks*' (the 'Project').

In this Deed **Confidential Information** includes all GPS track data including location, time and depth information given by the Trawl Fisherman to DPI and its officers in the course of obtaining track data for the Project.

By this Deed DPI agrees during and after the termination of this Deed to:

- Use the Confidential Information only for the purposes of the Project and for dissemination of the results from the Project;
- If the data is to be used for scientific forums or publications, present the Confidential Information in such a manner that the individuals, tracks or areas can not be identified;
- Except as provided above, maintain the confidentiality of that Confidential Information until it comes into the public domain (other than through a breach of this Deed), and not to use that Confidential Information;
- Store all documents or thing containing or embodying any part of that Confidential Information in its possession, custody or control securely;
- If DPI is uncertain of the status of any information, document or thing; to treat that information, document or thing containing or embodying information as containing or embodying Confidential Information of the Trawl Fisherman, and
- Immediately notify the Trawl Fisherman if it becomes aware of any breach or potential breach of this Deed.

Executed as a Deed

Signed sealed and delivered by
Dr Warren Hoey, Director-General

…..……………………………………………………...…..

Date: …..………………………………………..

Signature of Witness: …..………………………………………..

Name of Witness (Print): …..………………………………………..

18.4 Appendix D: TrackMapper Installation and User Guide

TrackMapper 1.0c

Installation and User Guide

20 Dec 2005

Norm Good

Installation

Basic program

The program executable 'TrackMapper.exe' is supplied along with two folders, called 'input' and 'output'. The input directory contains files required for the program to look up parameter values such as database connections, fishery definitions and other miscellaneous information which will be discussed further in this document. The output directory will contain all output from TrackMapper consisting of ArcView shapefiles. The program and folders can be placed in any folder on your machine. You can also create a shortcut to your desktop if desired.

Importing VMS_database Oracle table

It is assumed that you have a copy of Oracle on your machine with a database suitable for importing the vms_database table. A copy of this table is supplied under the Oracle folder on the CD supplied.

To import…

C:\> IMP <database name>/<password> file=<CD drive letter>:\VMS_DATABASE.DMP

This will create the table VMS_DATABASE in your database.

Creating an ODBC link from TrackMapper to Oracle

First go to Start menu > Programs > Administrative Tools > Data Sources (ODBC) > System DSN tab.

Ensure that a Data Source is named for your Oracle database.

- For Oracle 9i®, use 'Microsoft ODBC for ORACLE' driver

To create a New Data Source click the Add button. Select the 'Microsoft ODBC for ORACLE' driver and click 'Finish'. Type in 'database' in the 'Data Source Name' window and click 'OK'.

TrackMapper

What TrackMapper does

TrackMapper is used to:

- Produce effort maps at any spatial scale
- Produce catch and CPUE maps at any spatial scale
- Produce individual tow maps.

Running TrackMapper

For creating standard maps TrackMapper can be run straight from the console by running TrackMapper.exe. You will then be asked a series of prompts before the main data program executes. A sample console input is shown in the following screen output.

```
"C:\MERMAED_VMSUnit\TrackMapper\Debug\GetData.exe"

    ******************************************************
    *                                                    *
    *                  VMS TrackMapper                   *
    *                                                    *
    *                    Version 1.0c                    *
    *                                                    *
    *          Copyright Good and Peel 2005 QDPI         *
    ******************************************************


 Do you wish to produce a Catch or CPUE map (y/n)
n

 Do you wish to produce a map of trawl effort y/n)
y
 Do you wish to produce map of trawl paths (y/n)
y
 Do you wish to look at:
   1- Targeted Catch Only
   2- Any catch instance
   3- Top three catch
2
 Minimum longitude decimal (e.g. 152.25)
151

 Maximum longitude decimal (e.g. 154.5)
154

 Minimum latitude decimal (e.g. -28)
-29

 Maximum latitude decimal (e.g. -21.3)
-10

 Enter start date with a space between dd mm yyyy (e.g 01 01 2004)
01 01 2002

 Enter end date with a space between dd mm yyyy (e.g 01 02 2004)
01 10 2002

  Please enter gridsize
  1- one minute
  6 - CFISH GridSite
  30 - CFISH Grid
  or any other size you wish to plot

1.5

  This section asks you how many fisheries you want to
  map. Normally you would require information for specific
  fisheries such as EKP, Banana, or Scallop. However you have the
  option to combine data from several fisheries.

  Please enter how many fisheries to map, e.g. (EKP and Scallop) or EKP
3

  Fishery codes are
  EKP (Eastern King Prawn)
  Scallop (Saucer Scallop)
  TigerEndev (Tiger and Endeavour Prawns)
  Banana (Daytime banana prawn)

  All (all tracks irrespective of speed or time of day)

  Default (uses a general speed and time of day rule
  applied to all species.)

  Enter each fishery code followed by pressing the Enter key
EKP
Scallop
Default
```

The prompts are mostly self-explanatory, with the exception of a few points.

1. If you select 'y' to produce a map of catch and CPUE, effort will automatically be calculated and you will not receive the prompt 'Do you wish to produce a map of trawl effort'.
2. Targeted catch refers specifically to the first catch column in the CFISH logbook, irrespective of weight. For example, if you are mapping EKP targeted catch, any EKP catch not in column one will not be included in calculations.
3. Ensure that all latitudes are prefixed with the negative sign.
4. When selecting multiple fisheries to map ensure that you enter the same number of fishery codes (e.g. 'number of fisheries'=3, 'fishery codes'= EKP, Scallop, Banana). Note that the program is case sensitive so enter the fishery codes as shown.
5. After the last parameter has been entered the program will show the following messages;

   'Connected to VMS database'
   'Buffering more VMS data'

   The program will then scroll through the number of records accessed on screen, 10000 at a time. When TrackMapper has finished processing all records the program will terminate and the window will close. If the program exits before you see any records scrolling, an error may have occurred, possibly due to an error in user input.

The parameter inputs are the most common used to make a map. However, there are also text files in the 'input' folder which can be modified. Some of these are largely redundant and are only included for use in future version changes.

Input files

*.fsh files

Fisheries and the trawl signature definitions are contained in *.fsh files in the ‘input’ folder. For example the EKP fishery definition file when opened in Notepad contains the following:

```
**********************************************************************
EKP
This is the EKP Fishery
F
7
701303  701304  701305  701399  701904  701910  701917
149     156
-29     -21
17 0
7 0
0.1     2.3
0
```

Notes for the above

1. Name of fishery
2. Any notes
3. Flag to invert species selection below
4. Number of species
5. List of species
6. min long, max long for fishery
7. min lat max lat for fishery
8. start time, dawn/dusk (1 for dawn and 2 for dusk)
9. end time, dawn/dusk (1 for dawn and 2 for dusk)

if using dawn/dusk option then start/end time is time before (-ve) or after (+ve) relative time

10. min speed, max speed
11. number of boats with fishery based signatures
12. list of boat mobile numbers and their max speed

You can change any of the parameters in this file to suit your needs. For example if you wished to map all trawl NOT containing EKP then you would change the line 3 to ‘T’. You might like to change the name of the fishery to NOTEKP and save the file as NOTEKP.fsh and then select this when prompted in the console application. If you did not want to include species code 701917 in the list of species then change line 4 to ‘6’ and delete ‘701917’ from line 5.

There are two options for setting start and end times. The first is a set start and end time such as 5pm and 7am respectively. Alternatively you can set the times to coincide with dusk and dawn. For example, if you want to define a start time to be 2 hours before dawn then enter the following on line 8 (start time), -2 1. For 2 hours after dusk enter the following on line 9 (end time), 2 2.

A new algorithm is being developed for individual boat-based signatures so the list has not been included in this version of TrackMapper.

Anchorages.anc
This file contains the lat/longs of all anchorages determined by another program. Generally this file would be left unchanged. The anchorage buffer parameter in the settings.txt file determine the radius around each anchorage to exclude polls.

Fisheries.txt
This is a redundant file as the parameters are called from the console application.

Mapextent.txt
Similarly this file is largely redundant except for line 5 stating the buffer size for maps. The default is 0.5 degrees.
If you want to map the catch and effort from an individual boat then type in the boat's Mobile_no on line 9. This version only allows one boat at a time to be mapped but if required this can be a multiple boat option in future versions.

Maptype.txt
All the parameters required in this file are include in the console application.

Settings.txt
This file contains the following parameters:

****

0.01
1.25
6
1
100000

Notes
1. Anchorage buffer (in degrees)
2. Maximum time between polls (hours)
3. Minimum number of boats to count in each grid (mainly a space saving option)
4. Grid size in minutes (do not change, include in console application)
5. Maximum Entropy size (do not change, for making a resource map)

The user is free to change the settings of these parameters except gridsize which determined by user input and maximum entropy size which has been disabled.

Note: Increasing the minimum number of boats on line 3 increases array sizes considerably. If this number is too large the application may run out of memory and cause a run-time error.

Databaseinfo.txt

This file gives specific information for the database which should not need to be changed once the software is setup. The lines in the file are:

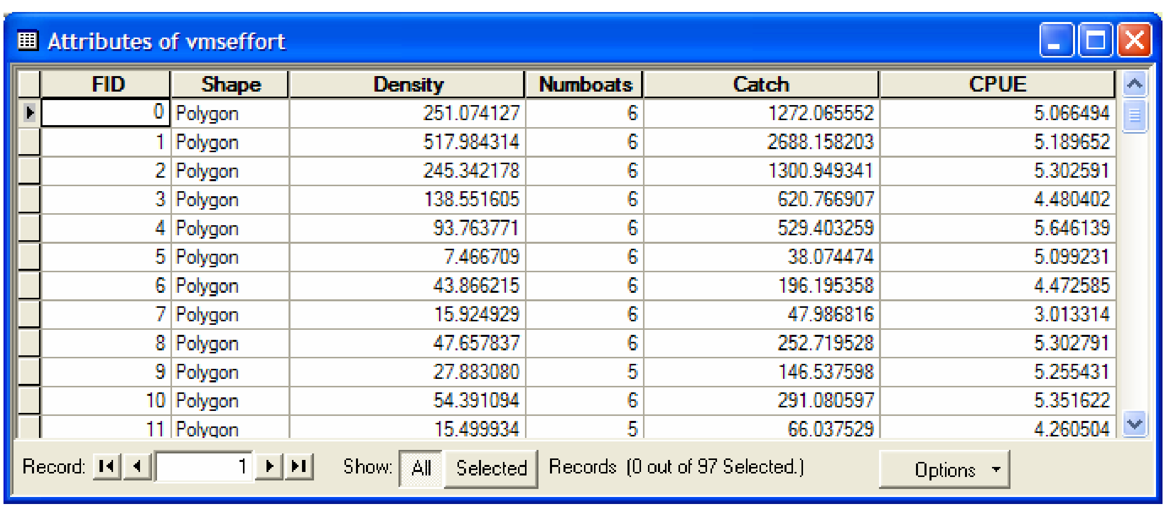
name of ODBC link
database user name
database user password
name of VMS table

an example is as given below

Database
TerraVision
TerraVision
VMS_DATABASE

Output files

For catch, CPUE and effort there are polygon files in the form of grids containing aggregate information as shown in Figure 18.1. The files associated are named vmseffort.dbf, vmseffort.shp and vmseffort.shx.



Attributes of vmseffort

FID	Shape	Density	Numboats	Catch	CPUE
0	Polygon	251.074127	6	1272.065552	5.066494
1	Polygon	517.984314	6	2688.158203	5.189652
2	Polygon	245.342178	6	1300.949341	5.302591
3	Polygon	138.551605	6	620.766907	4.480402
4	Polygon	93.763771	6	529.403259	5.646139
5	Polygon	7.466709	6	38.074474	5.099231
6	Polygon	43.866215	6	196.195358	4.472585
7	Polygon	15.924929	6	47.986816	3.013314
8	Polygon	47.657837	6	252.719528	5.302791
9	Polygon	27.883080	5	146.537598	5.255431
10	Polygon	54.391094	6	291.080597	5.351622
11	Polygon	15.499934	5	66.037529	4.260504

Record: 1 Show: All Selected Records (0 out of 97 Selected.) Options

Figure 18.1 Vmseffort shape file attribute table (example only, no real data used).

For maps of individual tows the output is a polyline shapefile containing information including Mobile_no (example only), Date, Time of day, Polling interval, Calculated speed, Start longitude, Start latitude, Catch1, Species1,…Catch6, Species6 (Figure 18.2). The files associated are named vmstows.dbf, vmstows.shp and vmstows.shx.

Attributes of vmstows

FID	Shape	Mobile_no	Date	Time_Day	Polltime	CalcSpeed	slongitude	slatitude	Catch1	Species1	Catch2	Species2
0	Polyline	1	4/01/2002	19.9	1	1.0611	153.19	-26.708	50	701304	0	0
1	Polyline	1	4/01/2002	20.9	1	1.0962	153.2073	-26.7387	50	701304	0	0
2	Polyline	1	4/01/2002	21.9333	1.0333	0.3388	153.226	-26.77	50	701304	0	0
3	Polyline	1	4/01/2002	22.9333	1	0.5148	153.226	-26.7813	50	701304	0	0
4	Polyline	1	4/01/2002	23.9333	1	0.1933	153.226	-26.798	50	701304	0	0
5	Polyline	1	5/01/2002	0.9333	1	1.0722	153.2307	-26.8027	50	701304	0	0
6	Polyline	1	5/01/2002	1.9667	1.0333	1.0617	153.2287	-26.768	50	701304	0	0
7	Polyline	1	5/01/2002	2.9667	1	0.343	153.22	-26.7333	50	701304	0	0
8	Polyline	1	5/01/2002	3.9667	1	1.0902	153.2107	-26.726	50	701304	0	0
9	Polyline	1	5/01/2002	4.9667	1	0.4824	153.2273	-26.758	50	701304	0	0
10	Polyline	1	5/01/2002	5.9667	1	2.8101	153.2307	-26.7733	50	701304	0	0

Record: 0 Show: All Selected Records (0 out of 8480 Selected.) Options

Figure 18.2. Vmstows shape file attribute table (example only, no real data used).

Additional notes

The ability to produce resource maps using Maximum Entropy modelling has been disabled in this version. Resource maps are used mainly for research and stock assessment purposes and model output needs to be interpreted correctly before maps are produced.

18.5 Appendix E: Potential application of VMS data for the identification of shallow and deep trawling areas in the EKP fishery (Chapter 10)

A map of all fishing effort from Dec 2000 to July 2004 at the one-minute scale was made and divided into shallow and deep areas using a 50 fathom (90 m) contour polygon (which defines the cut-off between deep and shallow sectors). For those 30-minute grids that were intersected by the 50 fathom contour, the amount of trawl effort (in hours) defined as shallow was divided by the total amount of trawl effort within the grid. An extract from the resulting table is shown below:

Grid	Total Effort	Shallow effort	Deep effort	Proportion Shallow	Proportion Deep
U30	82983.09	3920.87	79062.22	0.05	0.95
U31	3997.32	3997.32	0.00	1.00	0.00
U32	14648.57	14549.74	98.82	0.99	0.01
U33	0.09	0.00	0.09	0.00	1.00

A visual version of the above table is shown in Figure 18.3. It shows the 50 fathom contour and the split of effort between the deep and shallow sectors within selected 30-minute grids.

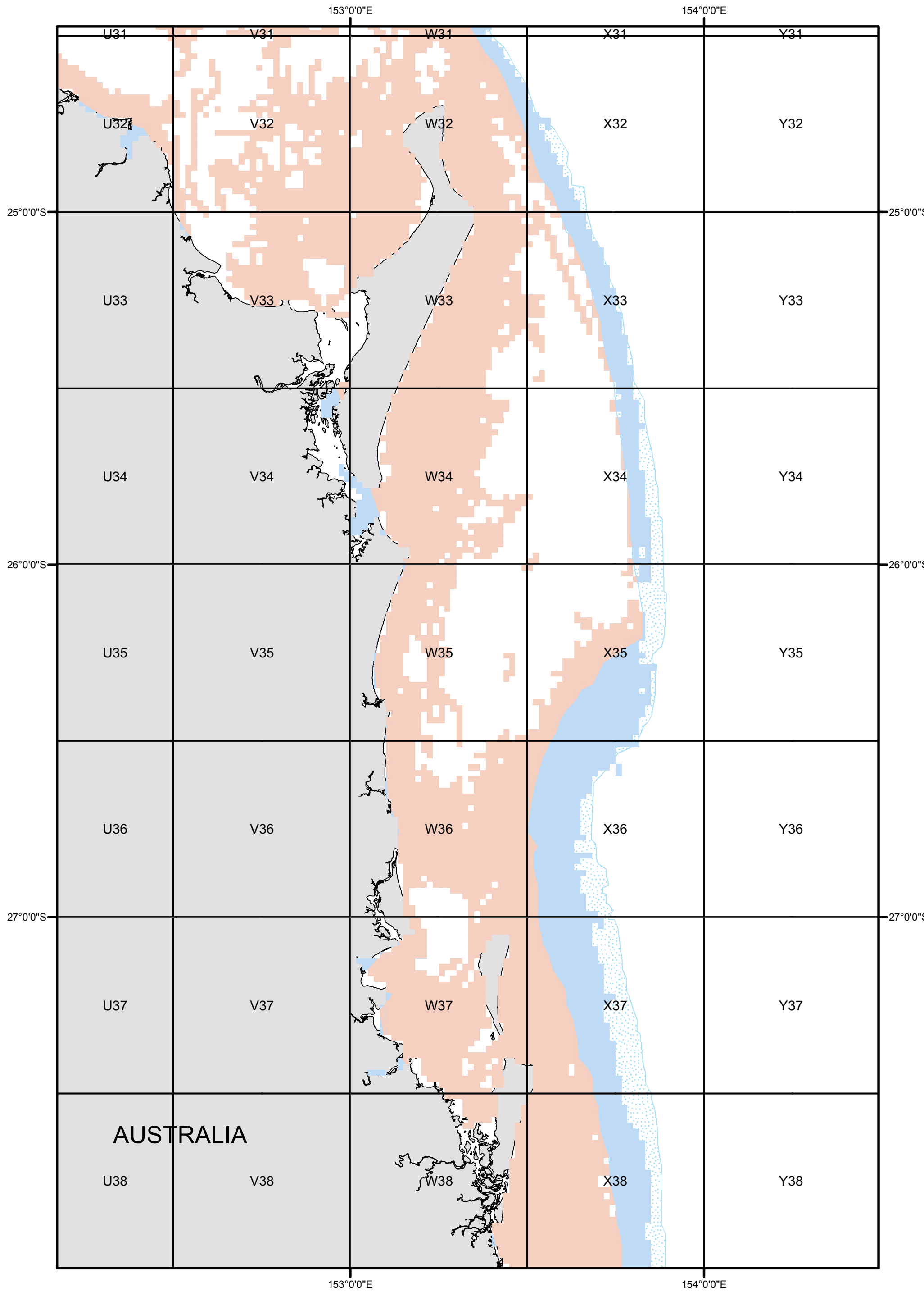


Figure 18.3. Total trawl effort from December 2000 to July 2004. Pink areas are in the shallow sector and blue areas are in the deep sector. Blue areas close to the coast are due to errors in the construction of the 0–50 fathom contour polygon.