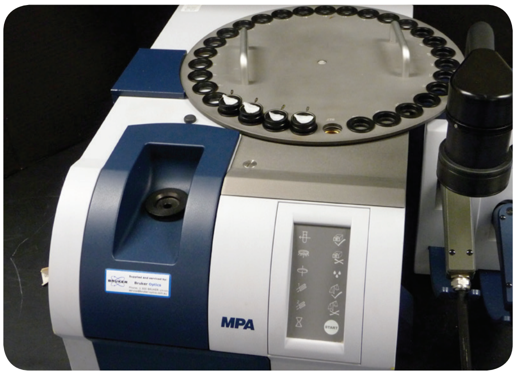
Department of Agriculture and Fisheries

Revolutionising Fish Ageing:

Using Near Infrared Spectroscopy to Age Fish



J.B. Robins, B.B. Wedding, C. Wright, S. Grauf, M. Sellin,
A. Fowler, T. Saunders and S. Newman
2015

FRDC Project No 2012/011

CS4052 04/15

FRDC
FISHERIES RESEARCH & DEVELOPMENT CORPORATI

Government of Western Australia
Department of Fisheries

Northern Territory Government

SARDI
SOUTH AUSTRALIAN RESEARCH AND DEVELOPMENT INSTITUTE

Queensland Government

The Fisheries Research and Development Corporation and The State of Queensland (through the Department of Agriculture and Fisheries) [2014]. All rights reserved.

ISBN 978 0 7345 0449 4

FRDC Project Number: 2010/011 Revolutionising Fish Ageing

Ownership of Intellectual property rights

Unless otherwise noted, copyright (and any other intellectual property rights, if any) in this publication is owned by the Fisheries Research and Development Corporation and The State of Queensland (through the Department of Agriculture and Fisheries).

This publication (and any information sourced from it) should be attributed to: Robins J.B., Wedding B.B., Wright C., Grauf S., Sellin M., Fowler A., Saunders T. and Newman S. (2015). Department of Agriculture, Fisheries and Forestry, Revolutionising Fish Ageing: Using Near Infrared Spectroscopy to Age Fish. Brisbane, April, 2015. CC BY 3.0

Creative Commons licence

All material in this publication is licensed under a Creative Commons Attribution 3.0 Australia Licence, save for content supplied by third parties, logos and the Commonwealth Coat of Arms.

Creative Commons Attribution 3.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. A summary of the licence terms is available from creativecommons.org/licenses/by/3.0/au/deed.en. The full licence terms are available from creativecommons.org/licenses/by/3.0/au/legalcode. Inquiries regarding the licence and any use of this document should be sent to: frdc@frdc.gov.au.

Disclaimer

The authors do not warrant that the information in this document is free from errors or omissions. The authors do not accept any form of liability, be it contractual, tortious, or otherwise, for the contents of this document or for any consequences arising from its use or any reliance placed upon it. The information, opinions and advice contained in this document may not relate, or be relevant, to a readers particular circumstances. Opinions expressed by the authors are the individual opinions expressed by those persons and are not necessarily those of the publisher, research provider or the FRDC.

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.

Researcher Contact Details

Name: Dr Julie Robins
Address Agri-Science Qld, DAF, Ecosciences Precinct
Phone: Joe Baker St, Dutton Park Q 4102
Fax: 07 32554230
Email: 07 38444529
julie.robins@daf.qld.gov.au

FRDC Contact Details

Address 25 Geils Court
Deakin ACT 2600
Phone: 02 6285 0400
Fax: 02 6285 0499
Email: frdc@frdc.com.au
Web: www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

Cover photos: (i) otoliths on the carousel of a Bruker Fourier-Transform NIR multi-purpose analyser (MPA); Snapper (Pagrus auratus); and (iii) otolith on the integrating sphere window of the multi-purpose analyser.

Contents

Tables

Figures

Acknowledgments

The following people assisted in developing an idea to a fully funded research project and their contributions are greatly acknowledged: Ian Halliday, Clive Turnbull, Sue Poole, Warwick Nash, Stuart Hyland, Chad Lunlow and Paul Exley. We also thank Sue Helmke, Jason McGilvray, Warwick Nash, Jonathan Staunton Smith, Wayne Sumpton, Stephen Wesche and Olivia Whybird for their patience in discussing the application of NIRS technology to fish ageing programs. We thank Quentin Allsop, David Fairclough, Sue Helmke, Bruce Jackson, Andrew Prosser, Steve Wesche and Olivia Whybird for their assistance with selecting and supplying fish otoliths to this project.

Abbreviations

ALK	Age-length-key
AP	Acquisition point
Aut	Autumn
cm^{-1}	Wavenumber
CV	Cross-Validation
Eqn.	Equation
FQ	Fisheries Queensland
GSV	Gulf St Vincent
LC	Length class
LR	Linear regression
LTMP	Long Term Monitoring Program
LV	Latent variables
MLR	Multiple linear regression
mth	Months
MPA	Multi purpose analyser
n	Number
NIR	Near Infrared
NIRS	Near Infrared Spectroscopy
OR	Outliers removed
OtoWt	Otolith weight
PLS	Partial least squares
QA	Quality assurance
Qld	Queensland
R^2	Coefficient of determination
R_c^2	Calibration model coefficient of determination
R_v^2	Validation coefficient of determination
RMSEC	Root mean square error of calibration
RMSECV	Root mean square error of cross validation
RMSEP	Root mean square error of prediction
RPD	Ratio performance deviation
SA	South Australia
SEC	Standard error of calibration
SECV	Standard error of cross validation
SEP	Standard error of prediction
SD	Standard deviation
SDR	Standard deviation ratio
Spr	Spring
Sum	Summer
WA	Western Australia
Win	Winter

Executive Summary

The current project is the first dedicated research project to apply Near Infrared Spectroscopy (NIRS) to age fish with structured input from fisheries agencies in Queensland, South Australia, Northern Territory and Western Australia. Results from the current ‘proof of concept’ study indicate that near infrared (NIR) spectra collected from fish otoliths have potential to estimate the age of Barramundi (*Lates calcarifer*) and Snapper (*Pagrus auratus*), with performance varying between species and locality of capture. A case study of hypothetical running costs suggest significant cost savings could be achieved if NIRS is used to supplement standard fish ageing methods. However, there is considerable time (i.e., at least 3 years) and start-up costs to develop and validate NIRS calibration models for fish age to a point where only model maintenance is required (i.e., running costs). Results also indicate that NIRS may be particularly useful for spatial (e.g. stock) discrimination. The potential applicability of NIRS was recognised by end-user stakeholders in Queensland and the Northern Territory, who are proposing further research work. Understanding what NIRS measures in fish otoliths and how this is correlated with age (or geographic location) was a common desire of fisheries end-users in all jurisdictions, because this knowledge could reduce error and would significantly enhance the applicability of NIRS technology in fisheries science.

Background

Ageing of fish is a central part of work by many fisheries agencies, with over 60,000 otoliths estimated to be collected and processed annually in Australia. The standard ageing method usually involves blocking otoliths in resin, cutting thin sections (~300 to 500 µm) using a diamond wafering blade, mounting sections on microscope slides, and viewing with either transmitted or reflected light under a microscope. This process takes about 15 minutes (over several days) per otolith, with the labour costs of this work making fish ageing a significant cost for fisheries agencies. Worldwide, there is a plethora of ongoing research into alternative methods to standard ageing that are faster, cheaper, automated and non-lethal.

NIR technology has been used for decades as a diagnostic tool in a wide range of science disciplines; primarily because it offers a rapid, repeatable and cost-effective method of predicting properties of interest e.g. moisture in wheat, oil content in sandalwood, ripeness or quality in fruit. The method relies on developing a calibration equation that relates the property of interest to NIR spectra measured by a spectrophotometer. The chemical component(s) in otoliths measured by NIR spectra that are correlated with age are at this stage unknown, but the correlation between the NIR spectra of otoliths and fish age were high in preliminary research on Saddletail Snapper (*Lutjanus malabaricus*).

Aims and objectives

The current project aimed to evaluate the innovative application of NIRS as a reliable, repeatable, and cost-effective method of ageing fish, using otoliths of Barramundi and Snapper as study species. Specific research questions included assessing how geographic and seasonal variation in otoliths affects NIRS predictive models of fish age, as well as how the NIR spectra of otoliths change in the short-term (i.e., <12 months) and long-term (i.e., historical otolith collections) and what effect this has on the predictive ability of NIRS models. The cost-effectiveness of using NIRS to supplement standard fish ageing methods was evaluated using a hypothetical case study of Barramundi.

Methods

Fisheries agencies from Queensland, South Australia, Northern Territory and Western Australia provided otoliths samples to the project from their standard fish length and age sampling programs. Otoliths collected in 2012 were considered ‘fresh’, where fresh is a relative term and refers to otoliths that are within the ‘normal time frame’ (i.e., usually weeks to months) that it would take an agency to process otoliths for standard age estimations. Otoliths collected from 2009, 2006 and 2003 were considered ‘historic’, having been stored by fisheries agencies for up to nine years after collection.

NIRS is a secondary method of determination, and requires calibration samples to develop a predictive model that is applied to other unknown samples. The predictive performance of any NIRS calibration model is dependent on the accuracy, error and bias inherent in the reference samples; in this case fish age based on visual assessment of thin otolith sections. Errors in the estimated visual age will perpetuate through the NIRS predictive models. In short, the more accurate and precise the visual age estimate, the more accurate and precise the NIR calibration and prediction. It is also important that the NIRS calibration set includes representative potential variability (e.g., across age range, length range, spatial or seasonal differences) so that the model is robust across the unknown predictive sample. However, estimates of biological fish age (in months) based on visual assessment of thin otolith sections have (unavoidable) associated error, because the absolute age of most wild fish is unknown. Absolute age cannot usually be estimated without further significant work (e.g., daily increment counts inside the first increment of the otolith to estimate the birthdate of each fish). The NIRS models developed in the current study were based on a generic set of wavebands and assumed that NIRS measures something within fish otoliths that is linearly related to fish age.

The efficacy of NIRS model outputs was assessed against several metrics that have relevance to fish ageing including: the percentage of a sample predicted within six to 12 months of observed age, percentage allocated to the correct age-class, presence of age-bias in predicted values, the Index of the Average Percent Error (IAPE) as a between-method measure of the standard fish ageing metric of 'precision' and statistical differences in age-class distribution.

Results

Results indicated that:

- NIR spectroscopy instruments could be readily configured to capture the NIR spectra of whole dry fish otoliths;
- NIR spectra were highly correlated with fish age in most instances, and
- NIRS calibration models had varying levels of predictive performance, in terms of R^2 (i.e., the extent to which the fitted straight line relationship explained the variability in the predicted fish age) and Root Mean Square Error (RMSE), which is a measure of the overall difference between observed and predicted values.

Performance was also variable in terms of the fisheries metrics that were used to assess efficacy.

In general, results suggested better NIRS estimates of Barramundi age from the Archer River (Queensland) than from the Fitzroy (Queensland) or Daly Rivers (Northern Territory); and better estimates of Snapper age from the Gulf St Vincent (South Australia) than from the Sunshine Coast (Queensland) or Mid-West Coast (Western Australia).

Geographic effects on the stability of NIRS calibration models were evident, being most prominent when models built for one location were used to predict fish age from another location. Geographic effects are not unusual in NIRS applications to agricultural produce (e.g., grains and fruits), where the biological properties of interest often vary geographically and seasonally. However, geographic effects could be accounted for in NIRS calibration models by including samples from all locations of interest and producing a robust model that had predictive performance (R^2 and RMSEP) comparable to geography specific models. Season (within year) had little effect on the predictive performance of NIRS models for Barramundi or Snapper, although seasonality should always be considered as a contributor to variability in NIRS.

We suspect that variability in results is probably a consequence of variable otolith microchemistry and how this differs between species, locations and year of collection, as well as differences in the collection, processing and storage procedures of otoliths (i.e., post-mortem handling).

The effects of storage on NIRS estimates of fish age were assessed by repeatedly acquiring NIR spectra from 'fresh' otoliths up to 17 months after collection. Results suggest that otolith chemistry stabilises somewhere between six and 11 months following collection for Barramundi and at about six months following collection for Snapper. The time difference in stabilisation may be the consequence of differences between species and their associated environmental conditions (e.g., freshwater/estuarine habitats compared to oceanic habitats), as well as differences in post-mortem handling of otoliths. Samples of ‘fresh’ and ‘historic’ otoliths analysed by the current project were insufficient to determine if otoliths degrade over storage times from one to more than five years, because of the confounding effect of between-year variability. Further work would be required to elucidate temporal variability from otolith degradation.

The current “proof of concept” study assumed a linear relationship between NIRS and fish age that was independent of growth rates. Improved models may be obtained through further research which could consider different data transformations prior to analysis and regional or seasonal specific NIRS wavelength selection.

Case studies of hypothetical running costs suggest significant cost savings could be achieved if NIRS is used to supplement standard fish ageing methods. However, there are considerable start-up costs over a number of years (i.e., three to five plus years) to develop and validate NIRS calibration models. It may be more feasible for fisheries agencies to outsource to dedicated NIR spectroscopy groups than to develop their own capability. Many fisheries agencies are part of large organisations that have expertise in NIRS servicing primary industries that could be tapped into, thus minimising capital expenditure costs, whilst retaining connection with the generation of fish age estimates.

NIRS based estimates of fish age may not be appropriate where the age structure is used to calculate recruitment indices or estimate year class strength, until issues around accuracy are better understood (see recommendations).

Implications

Standard ageing of fish otoliths via thin section is widely accepted as the current best estimate of observed age in Australia and overseas, despite its labour intensiveness, cost and limitations with accuracy and precision. The current ‘proof of concept’ study has made significant progress towards predicting fish age using NIRS collected from fish otoliths. Results indicated that fish age can be estimated through NIRS assessment of fish otoliths and potentially offers significant cost savings. However, further development is required to address issues with accuracy and precision before its use can be recommended for supplementing standard fish ageing methods in ongoing fisheries monitoring programs. Ultimately, each fishery agency will need to consider whether estimating fish age by NIRS is suitable for their species of interest and meets their data requirements. Results from the current work also indicated that NIRS may be a useful tool for spatial (e.g., stock) discrimination.

Recommendations

The application of NIRS technology to fish otoliths is innovative science worthy of further research.

The majority of fisheries collaborators in the current project wanted to know what NIRS measures in fish otoliths that is related to fish age. Understanding this relationship would assist in: identifying relevant wavelengths in the spectra and thereby minimising NIRS model error; identify the potential influence of growth rates on the relationship of otolith age and NIR spectra; thereby guiding post-mortem handling of otoliths for NIRS. This knowledge would lead to greater confidence in NIRS estimates of fish age and its suitability to supplement standard fish ageing methods. It should be noted that the question of what NIRS measures is never answered or fully confirmed in many industries that routinely apply NIRS. Ultimately, NIRS models need to provide predictions (of fish age) that are acceptable to, and meet the (data) requirements of (fisheries) end users.

Additional research identified from the current project (and discussed in detail in the main report) includes:

- Pilot application of NIRS to a complete 2012 sample of Barramundi age and length data from the southern Gulf of Carpentaria stock;
- Investigating the use of NIRS for spatial stock discrimination based on otolith microchemistry e.g., apply NIRS to the stock structure of tropical coastal reef fish in FRDC project 2013/017;

- Determining what NIRS measures in fish otoliths, how this related to fish age and implications for optimal post-mortem handling of otoliths for NIRS;
- Further repeat acquisition of NIR spectra from otoliths used in the current project to determine how otoliths change over the long term and impacts on the stability of NIRS calibration models;
- Application of NIRS to ‘difficult to age’ species such as small pelagic species, like pilchards; and
- Improving the performance of NIRS to age fish through more specific wavelength models.

Keywords: Near infrared spectroscopy, NIRS, chemometrics, fish ageing, otoliths, Barramundi (*Lates calcarifer)*, Snapper *(Pagrus auratus)*

Introduction

Background

This project was developed because preliminary data relating to the use of Near Infrared Spectroscopy (NIRS) to age Saddletail Snapper (*Lutjanus malabaricus)* found it was possible to construct NIRS predictive models which could predict the age of fish from their otoliths with a high degree of accuracy – R^2 of 0.94 and root mean square error of prediction of 1.54 increments[1] (Wedding *et al.* 2014). The concept of ageing fish using NIRS was discussed at the 2011 Workshop of the Australian Society of Fish Biology (ASFB), with keen interest in the concept expressed by the fish ageing sub-group, which contains representatives from all States and the Commonwealth. In Australia, about 60,000 otolith samples are collected annually, with an estimated cost of about $30 per sample for collection, preparation and visual assessment of thin otolith sections embedded in resin, hereafter referred to as 'standard' ageing methods. NIRS offers considerable scope to reduce the costs involved with ageing fish samples and removes some of the subjectivity associated with standard ageing methods. If successful, the adoption of NIRS would be a fast and reliable way of ageing large numbers of otoliths, potentially allowing management agencies to better target the expenditure of limited resources on fish ageing that underpin age-based fisheries management and stock assessment.

The preliminary results from the Saddletail Snapper project suggested that a rapid NIRS method of ageing fish would greatly reduce the cost of ageing per otolith, and possibly improve fish-ageing programs by:

- Allowing semi-automation of otolith age estimation for high sample throughput;
- Reducing delays between gathering of fish samples and the provision of age structures;
- Providing robust species-specific predictive models for ageing fish with calculated precision;
- Allowing greater flexibility to current ageing budgets;
- Transforming fish sampling protocols;
- Potentially leading to a national change to standardised fish-ageing procedures.

What is NIR Spectroscopy?

All organic matter is composed of molecules which consist of atoms; groups of which are linked together in various combinations mainly by covalent bonds. All molecules continually vibrate at specific frequencies. Irradiation of molecules by an energy source such as NIR light causes some molecules to change their vibrations from one energy level to another. When these transitions occur, energy is absorbed at a certain frequency coinciding with those of the molecular grouping in the scanned material. This absorption of energy is detected by NIRS instruments. Certain groups of small atoms, such as carbon-hydrogen (C-H), oxygen-hydrogen (O-H) and nitrogen-hydrogen (N-H), absorb at characteristic wavelengths. NIR spectroscopic measurements obtain information about the relative proportions of these fundamental absorbers which are also repeated throughout the NIR region as overtones or ripples of the fundamental absorber. Therefore, the chemistry of the otolith provides the spectral information that is assumed to be related to otolith age. If the chemistry changes throughout storage, then the NIRS calibration model will be affected. It is possible that: the amino acids within the otolith denature over time; pH may change through the loss of fluids and ions, as moisture levels stabilise; thus, resulting in a spectral change of the otolith from time of collection through subsequent

---

[1] The R^2 quoted for partial least squares regression models is the percentage of the total variance accounted for by the explained variance for the given number of latent variable in the NIRS model. It is not the square of the correlation coefficient. It is therefore possible to have a low R^2 but a high correlation coefficient suggesting, a strong linear relationship is present.

storage. Understanding otolith chemistry assists in understanding potential variability in NIR spectral information.

How do NIRS calibration models work?

The general NIR calibration process involves:

(i) reference or calibration sample selection of the property of interest - in this case fish age from otoliths;
(ii) evaluation of sample preparation and presentation for NIR analysis;
(iii) NIR spectrum measurement of reference otolith samples;
(iv) analysis of the otolith sample against the appropriate reference method, in this case standard sectioning and visual assessment. Note: the quality of the reference method value dictates that of the calibration model;
(v) chemometric model development (i.e., the calibration equation of the NIR spectra and chemical loadings combined mathematically to yield the calibration for analysis of unknown otolith samples);
(vi) validation of the calibration model to ensure that the model accurately predicts the property of interest (i.e., fish age) in otolith samples not subjected to the calibration process;
(vii) if the calibration model is found to be robust and accurate, the model (i.e., the relationship between visual age estimate and spectral data) can then be used to predict the age of further otoliths or new samples.

What does NIRS measure in fish otoliths that is related to age?

Spectral information is directly related to the chemical structure of the otoliths and is assumed to be related to otolith age. However, what specific chemical component(s) in the otolith that the NIR spectra correlate with, are at this stage unknown. Understanding otolith chemistry assists in understanding potential variability in NIR spectral information.

During otolith formation, calcium carbonate ($CaCO_3$) and protein are deposited on a daily basis, creating concentric increments that allow the age of the fish to be determined (Hale and Swearer 2008). Otolith form, size, weight, growth, consistency and chemical composition vary considerably among species (Zorica *et al.* 2010). Otolith structure displays alternate optically dense layers (rich in organic materials) and translucent layers (rich in minerals), forming concentric rings (Borelli *et al.* 2003) which are used in standard age estimation.

Otoliths consist of >90% (typically 90-96%) calcium carbonate by weight, 0.01-10% (typically 3-4%) organic matrix (protein complex) and approximately 1% non-organic trace elements (minor and trace elements, including radioisotopes and stable isotopes) (Campana 1999; Payan *et al.* 1999; Sohn *et al.* 2005; Chang and Geffen 2013). Otoliths are complex polycrystalline bodies composed of needle-shaped crystals of calcium carbonate radiating outwards in three dimensions from a centrally located nucleus and passing through a network of fibrous collagen-like protein called otolin (Marine-Institute 2007). Otolith calcium carbonate is in the form of twinned aragonite, although abnormal crystalline otoliths are composed of calcite or vaterite (Marine-Institute 2007; Parmentier *et al.* 2007). Otolin resembles keratin in its amino acid composition and is necessary as a “blueprint” in the mineralisation process (Marine-Institute 2007). Otolin contains water soluble proteins (acidic amino acids) and water-insoluble proteins (Marine-Institute 2007).

Biomineralisation of otoliths differs from that of vertebrate bone, molluscan shell and coral skeleton, since the otolith epithelium is not in direct contact with the region of calcification (Parmentier *et al.* 2007). The otolith is precipitated from the fluid of the endolymphatic sac of the inner ear (Radtke and Shafer 1992). Otolith growth is an acellular process, carried out away from the saccular epithelium, implying that the calcification process is strictly dependent on endolymph fluid chemistry (Borelli *et* *al.* 2003; Parmentier *et al.* 2007). The composition of endolymph is characterised by a high K^+ concentration, relatively low Na^+, an alkaline pH, saturated Ca^{2+} and HCO_3^- concentrations, has a low protein content, collagen and amino acids for otolith formation (Borelli *et al.* 2003; Parmentier *et al.* 2007). Calcium reaches the endolymph primarily from the blood plasma (Marine-Institute 2007). With the exception of Strontium (Sr), the minor elements are likely to be under physiological regulation via the endolymph (Swan *et al.* 2003).

Evaluation of otolith element concentrations reveals the interaction between biology and environment in the determining the otolith composition (Chang and Geffen 2013). The elemental composition of an otolith is influenced by physiological processes like metamorphosis, growth, age, food availability, activity levels, diurnal and seasonal cycles, reproductive status and environmental stress (Radtke and Shafer 1992; Tabouret *et al.* 2011). Species also vary in response to environmental conditions such as temperature and salinity. The structure and elemental composition of otoliths are influenced by many factors operating at many levels (Radtke and Shafer 1992; Tabouret *et al.* 2011). The most important factors influencing biomineralisation include:

- The source of mineral elements for the otolith, including trace elements, predominantly from the surrounding water (Chang and Geffen 2013);
- The uptake rate of different elements in the surrounding water which is influenced by fish metabolism, the method of entry into the fish (e.g., via gills or diet) (Marine-Institute 2007; Chang and Geffen 2013), salinity and temperature;
- The physical and chemical properties of calcium carbonate / aragonite which influence the incorporation of other elements into the otolith (Marine-Institute 2007);
- The chemical composition of water inhabited by the fish, which varies with geographic location, and depth (Marine-Institute 2007; Chang and Geffen 2013);
- The condition of the fish (e.g., gonad maturation, growth cycle, activity levels, age, stress levels) which will influence the composition of tissue and blood and affect the incorporation of different elements into the otolith (Marine-Institute 2007).

Otoliths are not chemically static structures, although they have great physical longevity, being found in middens (Owen 1998; Rowell *et al.* 2010). After removal from a fish, otoliths have a varied path of storage in different agencies. Some components of otoliths may alter in the short to medium term, with studies mostly focused on storage effects on elemental (trace metal) concentrations (Milton and Chenery 1998; Proctor and Thresher 1998; Rooker *et al.* 2001; Hedges *et al.* 2004). It is possible that the amino acids within the otolith will denature over time, pH may change through the loss of fluids and ions, and moisture levels will stabilise as the otolith dries out (Gauldie *et al.* 1998). As the chemistry of the otolith changes through storage, the spectral signal of the otolith may change from its time of collection and will probably affect the performance of the NIRS calibration model. However, we have limited information upon which to speculate how the performance of NIRS calibration models might be affected.

Need

Commonwealth and State legislation requires that fishery resources be managed sustainably. Age-based stock assessment methods are one of the more informative tools available for assessing Australia's fisheries. Collecting otoliths from recreational and commercial catches is undertaken across a wide range of fisheries as part of fishery monitoring and assessment programs Australia wide.

Estimating the age structure of fish populations is an important component of assessing the status of fished stocks, evaluating management strategies and assessing the impact of fishing. Fish age is usually determined by counting opaque bands (increments) in fish otoliths. This is a time-consuming process involving considerable preparation of the otoliths (i.e., resin-embedding, thin section cutting and sometimes polishing) prior to the visual increment count and edge assessment. The current project aimed to develop and validate a rapid innovative method for ageing fish based on an analysis of otoliths by NIRS. With increasing costs and, in some jurisdictions, diminishing research and development budgets, collection and ageing of representative fish samples becomes increasingly difficult. The use of NIRS on fish otoliths is a new application of established technology and if effective, potentially offers cost effectiveness that will be of global significance.

The project addressed the following Queensland Fisheries Research Advisory Board (QFRAB) priorities:

- Developing innovative tools and technologies for managing Australian fisheries;
- Developing more efficient, cost-effective ways of obtaining the information needed to undertake age-based fishery assessments;
- More reliable fishery assessments by improving the availability and quality of age information from fish population samples.

Objectives

1. Evaluate near infrared (NIR) spectroscopy as a reliable, repeatable, cost-effective method of ageing fish.
2. Determine the effect of geographic location (including latitude) distribution on NIRS algorithm stability.
3. Determine the effect (if any) of otolith storage time (years/months) on NIRS estimates of age.
4. Evaluate the cost-effectiveness of ageing fish by NIRS vs. standard otolith ageing, and develop optimised fish sampling regimes with respect to ‘cost’ (defined in terms of labour, lab time, field costs, etc.).

Methods

Objective 1. Evaluation of NIRS to age fish

The NIR calibration process, as outlined in the ‘Background’ section, was used to develop NIRS-based models for predicting fish age. Developing an appropriate calibration model requires reference or ‘training sets’ that cover not only the entire spectrum of quantities of interest (i.e., fish age), but also compositional space, instrument space and measurement condition space (e.g., sample handling and presentation). This avoids the need to extrapolate beyond the boundaries of the calibration set and makes the calibration robust and extensive. Temporal and spatial effects have major impacts on the robustness of the NIRS calibration models and must be incorporated into the development of the calibration model.

Reference otolith samples and age data

Two species of fish, Barramundi (*Lates calcarifer*) and Snapper (*Pagrus auratus*), were selected for study in the current project because of their extensive ranges in northern and southern Australia respectively. We investigated the relationship between age and NIR spectra for recently collected otoliths (referred to hereafter as ‘fresh’[2]) and for otoliths that had been stored for more than one year (referred to hereafter as ‘historic’) to determine if factors other than fish age influenced the calibration model.

Increment count and age data, based on the visual assessment of thinly sectioned otoliths embedded in resin, were supplied for each otolith sample by the respective state fisheries agency. Age, calculated in months (i.e., biological age), was a function of the date of collection, increment count (i.e., count of opaque bands), edge assessment and nominal birth dates for each species within a geographic location.

Barramundi

Fresh samples

Whole dry Barramundi otoliths were supplied by the Department of Agriculture and Fisheries (DAF) from two locations: (i) the Archer River estuary in the Gulf of Carpentaria (~13°30’S) that were collected in May 2012; and (ii) the Fitzroy River estuary on the Queensland east coast (latitude ~ 23°30’S) that were collected in February 2012 and September/October 2012 (Table 1). No fresh Barramundi samples could be obtained from the Northern Territory because there was no sampling for Daly River Barramundi in 2012.

Historic samples

Barramundi otoliths from archived DAF collections were used to explore the effect of long term storage (i.e., >12 months) on the ability to predict otolith age using NIRS (Table 1). Samples of Barramundi from the Daly River (Northern Territory) from 2011 were also used to develop an NIRS calibration model predicting fish age.

---

[2] ‘Fresh’ is a relative term and refers to the age of the sample since its collection. For otoliths, ‘fresh’ refers to otoliths that are within the ‘normal time frame’ that it would take an agency to process otoliths for age estimation – probably weeks to months, depending on the agency.

Snapper

Fresh samples

Whole dry Snapper otoliths were supplied by DAF from offshore waters of the Sunshine Coast, Queensland (~26°30’S), that were collected between June and August of 2012 (Table 1). Snapper otoliths were also supplied by the South Australian Research and Development Institute (SARDI) from the Northern and Southern Gulf St Vincent (NGSV and SGSV respectively) in South Australia (~35°S) that were collected between January and September 2012. No fresh Snapper samples were obtained from Western Australia.

Historic samples

Snapper otoliths from archived collections from the offshore waters of the Sunshine Coast (Queensland) and Gulf St Vincent (South Australia) were used to explore effects of long term storage on the ability of NIRS to predict otolith age (Table 1). Samples of Snapper from the Mid-West Coast of Western Australia (Fairclough *et al.* 2014) from 2010 were also used to develop an NIRS calibration model predicting fish age.

NIR Spectroscopy configurations and sample presentation

NIR spectroscopy systems can be operated in either diffuse reflectance, transmittance or interactance mode. The relative advantages and disadvantages of each mode are dependent largely on the physical characteristics of the sample of interest. Configuration trials compared NIRS systems operating over different wavelength ranges and identified which mode provided the most accurate and robust calibration model(s) across a test sample of Queensland Barramundi and Snapper otoliths.

Various sample presentation options were also assessed to determine a suitable method to present the otolith to the NIRS instrument to obtain the optimum spectral information. Once the optimised configuration, instrument and sample presentation method had been selected, spectra from all otolith samples were collected using the selected instrument configuration.

The NIRS instruments and configurations trialled to determine the optimal method for collection of the NIR spectrum from otoliths included different:

(a) Instruments:

- a Bruker Fourier-Transform NIR Multi-Purpose Analyser (MPA, Figure 1), with a carousel or integrating sphere (Figure 2);
- a Bruker Fourier-Transform NIR Matrix-F with an external emission head in reflectance mode; and
- a MicroPhazir NIR hand held unit.

-

Figure 1. **Bruker Fourier-Transform NIR Multi-Purpose Analyser (MPA), with a carousel**

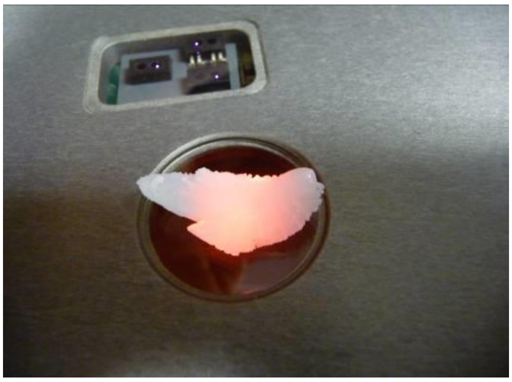


Figure 2. **Otolith on the integrating sphere window of Multi-Purpose Analyser (MPA)**

(b) ***NIR modes:***
Detection of NIR light from a sample may be by either transmission, reflection, transflection or interaction (Figure 3). In this instance, **reflectance** and **transmission** mode were assessed. In the case of transmission, incident light illuminates one side of the sample and the transmitted light may be detected from the other side. In the case of reflection, incident light illuminates the surface of the sample and the diffusely reflected light from the surface or from a portion near the surface may be detected.

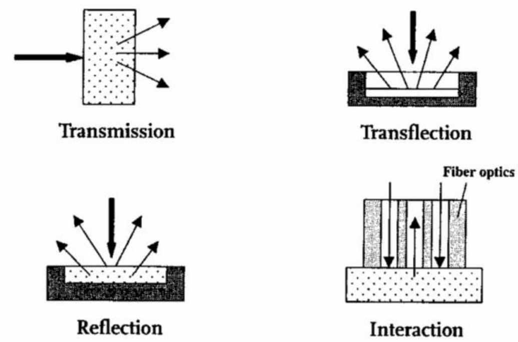


Figure 3. **Representation of transmission modes of NIR light, from Kawano (2002)**

(c) ***Spectra averaging:***
- 8, 16 and 32 (number of scans per samples).

(d) ***Otolith orientation:***
- concave and convex.

(e) ***Otolith presentation:***
- rotation through 0, 45, 90 and 180° during presentation on the MPA integrating sphere window.

(f) ***Otolith spectral differences:***
- between left and right side otoliths of the one individual.

(g) ***Operator differences:***
- for placement variation.

Spectra collection

Based on the findings from the above NIR configuration and sample presentation trials, NIR spectral data from all otoliths was collected on a Bruker Multi-Purpose Analyser (MPA) using an integrating sphere in diffuse reflectance mode, spectra averaging of 16 to 32 scans per second at a resolution of 8 cm^{-1}, with a concave up otolith orientation. The spectra collected from the fresh and historic otolith samples for Barramundi and Snapper are outlined in Table 1.

Table 1. Otolith samples used in the acquisition of NIR spectra

Species	Geographic Location	Nominal Birth Date	Collection Date(s)	Analysis & Comparison	NIR Spectra Acquisition Point [a] (Months)
Barramundi	Fitzroy River Estuary, Queensland (~23°30'S)	1 January	February 2012	Fresh	5, 6, 8, 11, 14, 17
			September & October 2012	Fresh Seasonal	
			2003, 2006, 2009	Historic	
	Archer River Estuary, Queensland (~13°30'S)	1 January	May 2012	Fresh	3, 4, 6, 9, 12, 15
			2006, 2009	Historic	
	Daly River Estuary, Northern Territory (~13°30'S)	1 January	July & September 2011	Historic	
Snapper	Sunshine Coast Offshore, Queensland (~26°30'S)	1 July	June to August 2012	Fresh	3, 6, 12, 15
			2006, 2009	Historic	
	Gulf St Vincent, South Australia (~35°S)	1 January	January to May 2012	Fresh	
			June to September 2012	Fresh Seasonal	
			2009	Historic	
	Mid-West Coast, Western Australia	1 August	January to November 2010	Historic	

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired. Three samples had repeat NIR spectra acquisition points. Further details are provided in the 'Results' section for Objective 2.

Data Analysis

Data analysis was carried out using "The Unscrambler" Software Version 9.8 (Camo, Oslo, Norway). Before the development of a calibration model, the variation of all spectral data was investigated by principal component analysis (PCA) and obvious atypical spectra were recorded as outliers and removed from further analysis. Partial Least Squares (PLS) regression was used to build the calibration models based on biological age in months of the diffuse reflectance spectral data. PLS regression attempts to establish a correlation between the spectral data and the visual assessment of otolith age in months (i.e., reference data set) to find the optimal model. Simply, the calibration equation of the NIR and chemical loadings combine mathematically to yield the calibration model which is then used for analysis of future unknown samples. Prior to PLS regression, raw spectral data were mathematically transformed to remove defects observed in the NIR spectra (e.g., noise and baseline drift). For all calibration models in the current study, the spectral data were transformed prior to model development using a combination of a 25-point Savitsky-Golay (SG) spectral smoothing (2nd order polynomial) and a first derivative transformation (25 point SG smoothing and 2nd order polynomial).

The objective of smoothing spectral data is to reduce noise, which can be described as random high-frequency perturbations. The Savitsky-Golay smoothing fits a low-degree polynomial through the data points within the local spectral window and derives the process signal values from the polynomial's function. With higher order polynomials, the individual weights derived from the polynomial coefficients are not the same for all data points within the spectral window to give a weighted moving average. Derivatives of spectral data are used to remove baseline shift and the resolution of overlapping peaks to enhance the visual resolution. Derivative mathematics provide a function of the slope change in the log reciprocal spectrum, which reduces the *y*-axis parallel shift caused by particle size and variation and normalizes the bases of the absorbance peaks, which lie on a slope across the spectral range. The first and the second derivatives are most commonly used, where the first derivative is the slope function of the log reciprocal spectrum and the second derivative is the slope function of the first derivative.

Full cross-validation (also known as jacknifing with segment size one) was used for small sample sets (<100 samples), while for larger sample sets (>100 samples) an independent test (validation) set was used. The validation set was obtained by randomly dividing the samples into a calibration set from which the calibration model was derived, and the remaining samples formed the validation (test) set which were then predicted by the calibration model. Cross-validation allows for the calculation of calibration statistics such as root mean square error of cross-validation (RMSECV) when only a small number of samples are available.

The assessment of model performance and robustness (ability to predict independent samples) from an NIRS perspective was based on the following partial least squares statistics:
(i) the coefficient of determination (R^2) of calibration (R_c^2) and validation/prediction (R_v^2);
(ii) the root mean square error of cross validation (RMSECV) and root mean square error of prediction (RMSEP);
(iii) the bias- (average difference between predicted and observed values;
(iv) the slope of the calibration/validation model; and
(v) the standard deviation ratio (SDR).
The SDR statistic is a measure of the ability of an NIRS model to predict a constituent and enables comparison of model performance across populations with different standard deviations (Baillères *et al.* 2002; Golic and Walsh 2006).

The R^2 quoted for PLS regression models is the percentage of the total variance accounted for by the explained variance for the given number of latent variables in the model. It is not the square of the correlation coefficient. It is therefore possible to have a low R^2 but a high correlation coefficient suggesting a strong linear relationship is still present but the predicted values are not close to the target one-to-one line. This can occur when there is large bias which is consistent across the full age range.

Efficacy of NIRS to predict fish age from fisheries perspective

The efficacy of ageing fish based on NIRS calibration/prediction models was assessed by comparing the performance measures between observed and predicted values for models of Barramundi ≤120 months old and Snapper <156 months old.

Efficacy from a fisheries perspective was viewed as the overall performance of predicted values across the following performance measures:

(i) the percentage of a sample set that was correctly allocated to its observed age-class/group;
(ii) the percentage of a sample whose predicted age was ±6 months of its observed age;
(iii) the percentage of a sample whose predicted age was ±12 months of its observed age;
(iv) the visual trends in age-bias plots of predicted values;
(v) the Index of Average Percent Error (IAPE), a standard fish ageing metric of 'precision; and
(vi) statistical differences in age-class distribution using a Kolmogorov-Smirnov test.

NIRS predicted age (in decimal years) was converted to age-class by rounding down, although rounding up, and rounding to the nearest integer were also examined for performance. It should be noted that some species have an age-allocation matrix to convert sample date, increment count and edge interpretation to an age-class (see Appendix 2). Age-class is the number of "birthdays" a fish is assumed to have had prior to its collection.

IAPE (Beamish and Fournier 1981) is a standard fish ageing metric of precision, where precision is defined as the reproducibility of repeated measurements on a given otolith, whether or not measurements are accurate (Chilton and Beamish 1982). It is traditionally based on the percentage agreement of repeat visual readings of otolith age in increment counts and is commonly calculated by some (but not all) of the collaborating fishery agencies as a part of the quality assurance of their fish ageing protocols. An IAPE value for repeat visual assessment of otolith age of <5% indicates sufficiently precise visual age estimates, while an IAPE of >10% indicates an unacceptable level of precision in the visual age estimates (Robertson and Morison 1999). In the current study, IAPE (presented in the performance measure tables) is used as a measure of the similarity of predicted age against observed visual age estimates i.e., a measure of between-method precision, compared to the standard use of IAPE as a measure of precision between readers using the same method.

Plots of observed versus predicted age and age-bias plots were developed for each species in each location. In the current study, age-bias plots provide a graphical measure of the age-by-age deviation of predicted (e.g., NIRS) age from the observed (i.e., visually assessed) age (Campana *et al.* 1995). Age predicted from NIRS models was then converted to age-class and plotted as an age-class distribution. Age-class distributions were compared using a Kolmogorov-Smirnov (K-S) test (at 0.05 level of significance) to determine if there were significant differences between the observed age distribution and the predicted age distribution.

Relationships between fish age and otolith weight were also developed for Barramundi and Snapper to provide a baseline against which to compare NIRS estimates of age. Otolith weight was selected because it is a simple and cost-effective parameter to collect and is highly correlated with age. Several studies have investigated the capacity to predict fish age from otoliths weight (Francis *et al.* 2005; Ochwada *et al.* 2008); some specific to the species of interest in the current project (McDougall 2004). However, otolith weights or other proxy methods are not widely used to estimate age mostly because of insufficient accuracy, despite significant research efforts to find an acceptable method (Worthington *et al.* 1995; Francis and Campana 2004; Lou *et al.* 2007; Lepak *et al.* 2012).

In the current study, the same calibration/validation or cross-validation otolith sets used in NIRS based predictions were analysed by regression to explore relationships between fish age, otolith weight and fish length, except where otoliths had major damage, resulting in their weight being an outlier. All regression models were developed using GenStat for Windows 16th Edition (Genstat 2013). The efficacy of results from the best regression model (highest adjusted R^2, lowest RMSE) was compared to results from NIRS using the fisheries performance measures mentioned above.

Objective 2. Effects of geographic and seasonal variation on NIRS

Barramundi

Barramundi otoliths sourced from two locations in 2012 (Fitzroy River and Archer River) were used to test for geographic differences and allow for the inclusion of geographic variability in calibration models. Samples from the Fitzroy River were sourced in February and September/October 2012, allowing seasonal variability to be included in model development and when combined with the Archer River samples, to include both geographic and seasonal variation.

Snapper

Snapper otolith samples obtained in 2012 were sourced from two locations (Offshore Sunshine Coast Queensland and Northern and Southern Gulf St Vincent in South Australia) and were used to test for geographic differences and allow for the inclusion of geographic variability in calibration models. Samples from the GSV had collections periods of January to May and June to September 2012, allowing seasonal variability to be included in NIRS model development for Snapper. Combining GSV samples with Queensland Snapper samples allowed for the inclusion for both geographic and seasonal variation.

Objective 3. Effects of otolith storage time on NIRS

Stored otolith samples may suffer slow degradation due to oxidation or other effects (e.g., moisture, mildew, light, heat exposure). These factors may result in spectral differences between otoliths and between the same otolith scanned at different NIR spectra acquisition points (i.e., storage time). These potential influences were assessed to determine whether the age-NIRS relationship of stored otolith samples could be used to estimate the age of future fresh samples.

Calibration models were developed for all NIR spectra acquisition points, that is, for all sample sets scanned at various times after collection (see Table 1). Preliminary analysis indicated improved model fit when otolith data sets were restricted to fish ≤120 months (10 years) of age for Barramundi and <156 months (13 years) of age for Snapper. Once each calibration model was developed and calibration statistics obtained, the individual models were then used to predict the biological age of each remaining sample set - at each NIR spectra acquisition point - for a species and location. The figures presented in the current report show the cross-validation/validation statistics and also the validation statistics obtained when a calibration model for a NIR spectra acquisition point is used to predict the ages of samples from a different NIR spectra acquisition point.

NIRS calibration models were developed using fresh otoliths of: (i) Barramundi from two locations (i.e., geographic variation) spanning up to 17 months after collection; and (ii) Snapper from one location spanning up to 18 months after collection.

NIR calibration models were developed using historic otoliths of: (i) Barramundi from two locations spanning 11 to 105 months in storage after collection; and (ii) Snapper from one location spanning 12 to 85 months after collection. These models were assessed for their variability in predicting otolith age, regardless of storage time.

Objective 4. Evaluate cost-effectiveness of ageing fish by NIRS

Cost-effectiveness

Each collaborating fisheries agency was asked to supply details on the collection and processing costs for the species they supplied to the current project. In addition, each collaborating agency was asked to describe how information collected on fish age was used by their agency to inform stock assessment and fisheries management.

The relative costs of ageing fish using traditional methods versus potential running costs using NIRS was compared using a case study of the 2012 southern Gulf of Carpentaria Barramundi stock.

Assumptions and Limitations

It should be noted that each state agency was asked to supply 10 otoliths from each age-class of each species from each location. In some instances (e.g., Queensland fresh otoliths), otolith samples were supplied for NIR spectra acquisition before standard ageing had occurred. Therefore, 10 otoliths from the main 20 mm length-classes were supplied in the assumption that this would provide at least 10 otoliths from the main age-classes in the various fisheries. Populations of Barramundi and Snapper (particularly from the Gulf St Vincent) also display variable recruitment and as such, sometimes it was not possible to supply 10 otoliths from year-classes when recruitment was very weak. As such, the otoliths analysed in the current project are indicative only of the samples supplied to the current project and cannot be assumed to be representative of the total populations of Barramundi and Snapper in their respective geographic locations.

The development of models to predict fish age based on NIRS assumes a relationship exists between what is measured by NIRS and fish age. As yet, it is uncertain as to what NIR spectra measures in relation to the chemical attributes of fish otoliths, and whether this is dependent solely on time or otolith growth or both.

It is important to note that NIRS is a secondary method of determination and relies on the accuracy of the reference method. The predictive performance of any NIRS calibration model is dependent on the accuracy, error and/or bias that is inherent to the reference samples, which in this case is fish age based on the visual assessment of otolith thin sections. Errors in the visual age assessment will perpetuate through the NIRS predictive models. In short, the more accurate the visual age estimate, the more accurate the NIR calibration and prediction.

Campana (2005) notes that ageing error includes: (i) "error that affects the accuracy or the closeness of the age estimate to the true value"; and (ii) "error that affects precision or the reproducibility of the repeated measurements on a given structure". The accuracy of visual age estimates of Barramundi and Snapper is based on the assumption that increments in these species are formed annually and thus relate to the age of a fish. It was generally agreed by the fisheries stakeholders assisting the project that annual increment formation had been validated for Barramundi and Snapper - for Barramundi, see Stuart and McKillup (2002), McDougall (2004) and Staunton-Smith *et al.* (2004); for Snapper, see Ferrell *et al*. (1992) and Francis *et al*. (1992).

The current project used increment count (in preliminary models) as well as estimates of biological age in months. Biological age is probably a more accurate estimate of true age than increment count, but it should be noted that biological age has an (unavoidable) associated error. Biological age is calculated based on the known capture date and increment count of an individual and an assumed nominal birth date for a species by stock or location. The nominal birthdate (i.e., 1st January for Barramundi and 1st July or August for Snapper) is unlikely to be the actual birth date (spawning date) for all individuals sampled. For Barramundi, the actual birth date is likely to be up to three months before or after the 1st of January; thus the true (or absolute) age of any individual Barramundi is potentially the estimated biological age ±3 months. A similar situation exists for Snapper. To remedy this "error" associated with estimated biological age from visually assessed thin sections of otoliths would require either: (i) counts of daily rings to estimate the absolute birth date of each individual within a calibration set; or (ii) fish from known spawning dates and therefore known absolute age.

Errors affecting the precision of the age estimates based on visual assessment of thin sections have been minimised as much as possible by obtaining age estimates from agencies with well-established ageing programs, some of which include a quality assurance (QA) approach, requiring training and qualification using reference collections of sectioned otoliths as recommended by Campana (2005). IAPE is generally considered as a metric of precision of visual age estimates. The IAPE's (of within method between-reader precision) for the Barramundi and Snapper samples considered in the current project were:

- 1.9% for 2012 Archer River Barramundi;
- 0.1% for 2012 Fitzroy River Barramundi (total sample, n~800);
- <0.1% for 2012 Fitzroy River Barramundi (February sample); and
- 0.7% for 2012 Sunshine Coast Snapper (total sample, n~600).

Results

Objective 1. Evaluation of NIRS to age fish – developing models

NIR spectroscopy configurations and sample presentation

Results of the configuration trials indicated that the best set up of the NIRS instruments trialled for whole dried otoliths was on the Bruker MPA using the integrating sphere in reflectance mode, with spectra averaging of 16 or 32 scans per sample at a resolution of 8 cm^{-1} and with an upward concave otolith (i.e., with the convex orientation exposed to the NIR light) at a 0° orientation (Figure 2). Otolith presentation (rotation from 0^o to 45^o, 90^o, 180^o in the concave orientation) on the integrating sphere window did not make a significant difference and there was not a discernable difference between operators for placement variation. Although, there was no significant difference found in this trial on the rotation of the concave otolith from 0^o, 45^o, 90^o, 180^o it is recommended that a consistent orientation be maintained to reduce possible discrepancies. This configuration was utilised for otolith sample NIR spectra collection throughout the project.

Results suggest there were no significant spectral differences between otoliths obtained from the left and right side, implying that either otolith of an individual can be used in the generation of NIRS calibration models (Figure 4). The outer edges of the box represent the 25th and 75th percentiles, with the 50th percentile (median) across the middle of the box. The whiskers represent 1.5 times the inter-quartile range, where the inter-quartile range equals the 75th percentile minus the 25th percentile. There are three data points outside this range with the outlying sample in red (at -3) being deformed. Fifty percent of the data lies between -0.33 and 0.27, and 95% of the data lies between -1.01 and 0.92 increment counts; thus indicating no systematic bias between left and right otoliths. In the current project, the left otolith was utilised in the majority of samples for developing NIR spectroscopy calibration models.

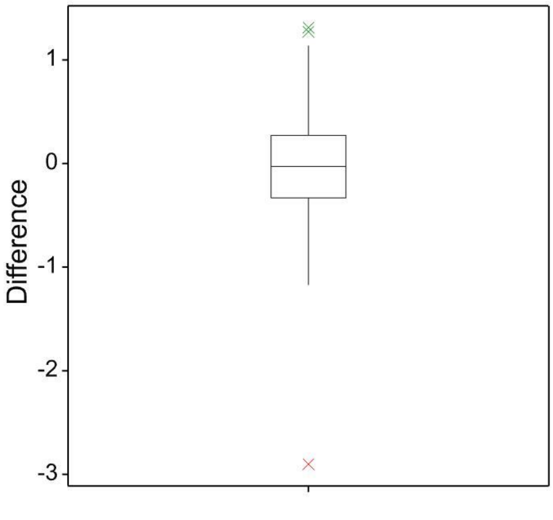


Figure 4. Box-and-whisker plot of the difference in estimated age (in increment count) from NIRS calibration models developed using left and right otoliths

Objective 1 (i.e., Evaluate NIRS as a reliable, repeatable, cost-effective method for ageing fish) is covered in the results within Objectives 2, 3 and 4.

Objective 2. Effects of geographic and seasonal variation on NIRS

Barramundi

Calibration model development

Relevant spectral information for ‘fresh’ Barramundi otolith calibration models was obtained primarily in the 4832 cm^{-1} to 4327 cm^{-1} NIR spectral region (waveband[3]). This relates to vibrational group frequencies related to -CH combination tones, H_2O and -OH combination tones, and –NH combination overtones. Carbonates contain strong NIR-active vibrational modes (Hunt 1977; Moron and Cozzolino 2003; Thomas *et al.* 2011). A typical absorbance spectrum for a whole Barramundi otolith is shown in Figure 5.

Calibration models were developed for each geographic location, season and a combined data set incorporating different locations using the generic wavelength selection based on the identified band assignments. Model development was trialled using estimates of both biological age (in months) and increment count. Biological age proved slightly more accurate than increment count (i.e., increased R^2 and decreased RMSEP). All models presented are for biological age in months.

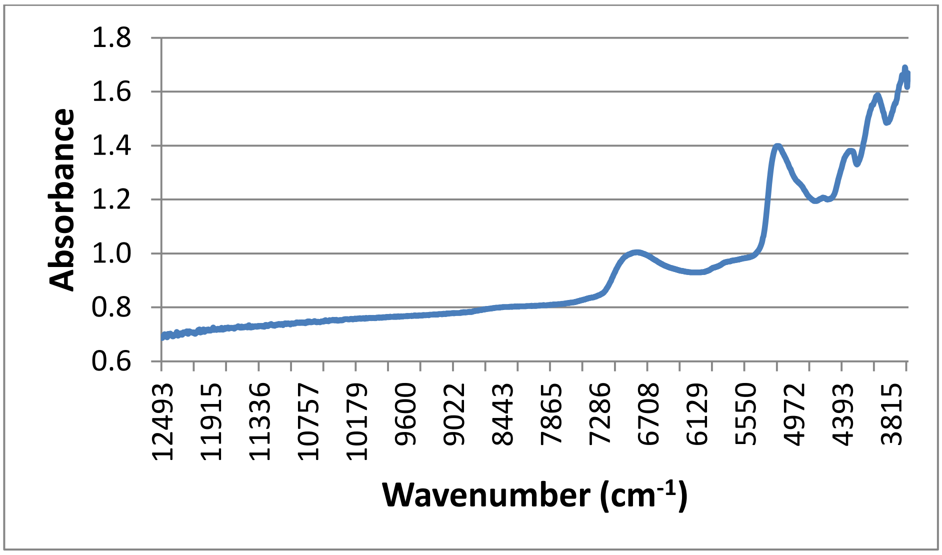


Figure 5. Typical raw absorbance spectrum for whole Barramundi otoliths

Geographic and seasonal variability based on generic wavelength selection

The PLS calibration and validation statistics for Barramundi otoliths from the Archer River (May 2012) and Fitzroy River (February and October 2012) based on a generic wavelength selection are presented in Table A2.1 (located in Appendix 3). The calibration statistics for the Archer River model were R_c^2 = 0.88 and RMSEC = 7.6 months and validation statistics were R_v^2 = 0.85, RMSEP = 6.6 months and an SDR = 2.61. The Fitzroy River calibration models had comparable predictive

---

[3] Waveband is a spectral region and represents a group of wavelengths. Wavenumber (cm^{-1}) and nm are a measure of wavelength.

performance, with an R_v^2 = 0.87, RMSEP = 7.5 months and an SDR = 2.80 for February 2012 samples; and R_v^2 = 0.90, RMSEP = 6.8 months and an SDR = 3.26 for October 2012 samples. The correlation coefficients (r_y) for the Archer River and Fitzroy River (February and October 2012) all suggest a strong linear relationship between the NIR predicted ages and the observed age for the validation sets (r_y = 0.94, 0.93, 0.95 respectively).

Archer River predicting Fitzroy River fish ages

The application of the Archer River (May 2012) calibration model to Fitzroy River samples collected in February and October 2012 (i.e., another location) was quite successful, but had reduced R^2's and SDR's, increased RMSEP's and substantial bias (Table A3.1). The predictive performance of the Archer River (May 2012) calibration model to the Fitzroy River February 2012 samples resulted in an R_v^2 = 0.63, SDR = 1.65, RMSEP = 14.1 months and a bias = -10.4 months; while those for Fitzroy River October 2012 samples were R_v^2 = 0.74, SDR = 1.98, RMSEP = 11.2 months and bias = -7.6 months. The large negative bias (i.e., -10.4 and -7.6 months) suggests that NIRS-predicted age of the samples in the validation set tended to be underestimated. These results were substantially improved through applying a bias correction, with a bias-corrected R_v^2 = 0.83 and RMSEP = 9.5 months and a bias-corrected R_v^2 = 0.86 and RMSEP = 8.3 months for February and October samples, respectively (Table A3.1). Although the R^2's decreased when a calibration model was applied to a different region, the r_y's were still ≥0.92 suggesting a strong linear relationship was present (Table A3.1).

Fitzroy River predicting Archer River fish ages

Similarly, the Fitzroy River (February 2012 and October 2012) calibration models were used to predict biological age of samples collected from the Archer River (May 2012). The Fitzroy calibration models had a reduced predictive performance for Archer River validation samples, with an R_v^2 = 'not reportable', RMSEP = 27.1 months and a bias = 25.6 months (Fitzroy February 2012 calibration model predicting Archer River 2012 validation samples) and R_v^2 = 0.56, RMSEP = 13.2 months and a bias = 10.7 months (Fitzroy October 2012 calibration model predicting Archer River 2012 validation set; Table A3.1). Application of bias correction improved the predictive performance to an R_v^2 = 0.80 and RMSEP = 8.9 months (Fitzroy February model applied to Archer validation set) and R_v^2 = 0.85 and RMSEP = 7.7 months (Fitzroy October calibration model applied to Archer validation set). The r_y's = 0.93 for both the Fitzroy River calibration models applied to the Archer River sample set.

Combined Archer and Fitzroy River samples

A combined geographic calibration model was developed using spectra from both the Archer River (May 2012) and Fitzroy River (February 2012) samples (n=198). This calibration model was used to predict the remaining Archer River (May 2012) and Fitzroy River (February 2012) samples (n=190). The predictive performance was comparable to models for individual geographic locations, with an R_v^2 = 0.86 and RMSEP = 7.6 months (Table A3.1).

Similarly, a combined geographic calibration model was developed using spectra from the Archer River (May 2012) and Fitzroy River (February and October 2012) samples (n=298). This calibration model was used to predict the remaining Archer River (May 2012) and Fitzroy River (February 2012 and October 2012) samples (n=289). The predictive performance was comparable to models for individual locations, with an R_v^2 = 0.87 and RMSEP = 7.8 months (Table A3.1).

These results indicate that calibration models for individual geographic locations were robust in terms of predictive performance (R^2 ≥0.85, RMSEP = 6.6 to 7.5 months, SDR's = 2.61 to 3.26). As with horticultural commodities, the application of a calibration model from one geographic location to another (i.e., Fitzroy River calibration model used to predict Archer River samples) was not successful (i.e., lower $R^2$and SDR, large RMSEP and large bias). However, the observed geographical variability may be the result of otolith microchemistry varying between locations. Geographical variability can often be addressed simply through bias correction. There was no apparent seasonal effect on the predictive performance of the Fitzroy River models (February calibration model applied to October 2012 validation set and vice versa, Table A3.1), although there was a tendency for reasonable bias (e.g., 4.8 and 3.9 months respectively). A combined geographic model incorporating biological variability from each location produced a robust model with a predictive performance comparable to individual geographic location models.

Model refinement

Otolith microchemistry is known to be affected by many factors (Radtke and Shafer 1992; Tabouret *et al.* 2011). Variability in otolith microchemistry is likely to be reflected in the NIR spectra. In the models discussed previously, Barramundi otoliths visually assessed as >120 months (10 years) were often under-estimated by the predictive model (Figure 6). With this in mind and given that the older Barramundi (>10 years) were only a small percentage of the total samples assessed by NIRS in the current project, we explored the applicability of models for Barramundi ≤10 years of age and for fish >10 years of age compared to models using all ages.

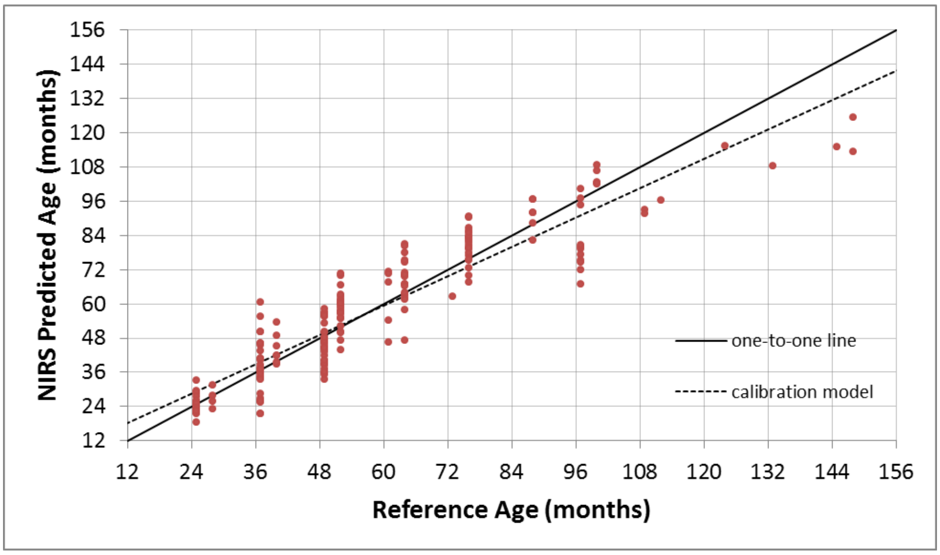


Figure 6. Observed age (reference age) versus NIRS predicted age of Barramundi all ages based on a generic wavelength calibration model using otoliths collected in 2012 from the Archer and Fitzroy Rivers all months combined (n=289, validation set)

The PLS calibration and prediction statistics for Barramundi otoliths for fish ≤120 months from both the Archer River (May 2012) and Fitzroy River (February and October 2012) based on the generic wavelength selection are presented in Table A3.2. There was not a sufficient number of samples in the current project to develop a calibration model for Barramundi >120 months. Calibration models based on Barramundi ≤120 months old resulted in slightly increased predictive performance i.e., similar R^2's, but with reduced RMSE's (Table A3.2), as compared to the models incorporating all ages of Barramundi (Table A3.1).

Geographic variability based on geographic specific wavelength selection

As well as restricting the age of visually assessed Barramundi otoliths to ≤120 months, the predictive performance of a calibration model can be marginally improved for each location through the selection of specific wavelengths for that geographic location (Table A3.3 *cf* Table A3.1). The specific wavelengths are most likely driven by the spatial variance in otolith microchemistry, in comparison to the application of a generic wavelength model (Table A3.1). As a geographic specific wavelength selection is designed for each location, the model's predictive performance is not as good as that of the generic wavelength selection model when applied to samples from other geographic locations (Table A3.3).

Snapper

Calibration model development

Relevant spectral information for ‘fresh’ Snapper otolith calibration models was obtained primarily in two main NIR spectral regions (wavebands) between 6160 and 4580 cm^{-1}. These spectral regions are similar to that found for Barramundi otoliths, and correspond to group frequencies related to –CH first overtones, C=C first overtone and –CONH$_2$ primary and secondary amides combination spectral regions. A typical absorbance spectrum for a whole Snapper otolith is shown in Figure 7.

Calibration models were developed for Snapper from each geographic location, season and a combined data set incorporating different locations using the generic wavelength selection identified for Snapper. Model development was trialled on both biological age (in months) and increment counts, with biological age proving slightly more accurate than increment count (i.e., higher R^2 and lower RMSEP). All models presented are for biological age in months.

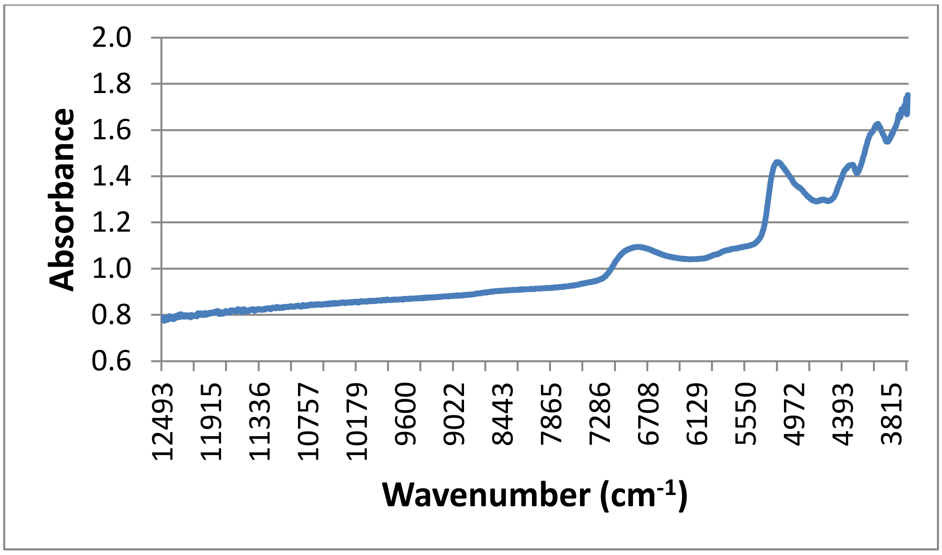


Figure 7. Typical raw absorbance spectrum for whole Snapper otoliths

Geographic and seasonal variability based on generic wavelength selection

The PLS calibration and prediction statistics for ‘fresh’ Snapper otoliths from the Sunshine Coast (Winter & Spring 2012) and Gulf St Vincent (Summer & Autumn 2012 and Winter & Spring 2012) based on a generic wavelength selection are presented in Table A3.4.

Validation statistics for Snapper from the Sunshine Coast indicated good model performance, with an $R_v^2 = 0.79$, RMSEP = 19.0 months and an SDR = 2.19. Similarly, the GSV Summer & Autumn 2012 samples had good predictive performance, with an $R_v^2 = 0.91$, RMSEP = 14.8 months and an SDR = 3.24 (Table A3.4). The Winter & Spring 2012 samples from the GSV (n=93) were evaluated using cross-validation and had high model performance, with an $R_{cv}^2 = 0.93$, RMSECV = 15.2 months and an SDR = 3.64 (Table A3.4). With small sample sets, there is a predisposition to select samples that represent as full a range as possible in the test set. Prediction of these tends to “flatter” the calibration model (Williams 2008), hence cross-validation is preferred over the validation set method. The r_y’s for individual locations were ≥0.89.

As seen with Barramundi, the calibration model for a specific geographic location and season was not as successful at predicting a different geographic location and/or season. For this reason, Table A3.4 only shows the calibration and prediction statistics for calibration models predicting on the same geographic location and season.

A combined calibration model was developed using spectra collected from the GSV over nine months of 2012 (i.e., January to September). The predictive performance of the combined seasonal model for GSV Snapper was comparable to results for the individual seasonal models for the GSV, with an R_v^2 = 0.89, RMSEP = 15.8 months, an SDR = 3.08 and r_y = 0.95, suggesting a strong linear relationship.

A combined geographic and seasonal calibration model was developed using spectra from both the Sunshine Coast (Winter & Spring 2012) and GSV (Summer & Autumn 2012 and Winter & Spring 2012) samples. The predictive performance was comparable to the individual GSV (Summer & Autumn 2012 and Winter & Spring 2012) models with a slight increase in predictive error, but better than the Sunshine Coast (Winter & Spring 2012) model, with an R_v^2 = 0.85, RMSEP = 17.9 months, an SDR = 2.61 and r_y = 0.92. These results indicate that individual geographic location calibration models were more robust in terms of predictive performance (i.e., $R^2 \geq 0.79$, RMSEP = 14.8 - 19.0 months).

In general, predictive performance (R_v^2, RMSEP, and SDR) was better for Gulf St Vincent Snapper than for Sunshine Coast Snapper. The reasons for this are unknown, but are possibly a consequence of the within-sample variability in otolith microchemistry or visual readability.

As with Barramundi, Snapper models had a tendency for otoliths ≥13 years to have underestimated NIRS predictions of age (Figure 8). Therefore, we explored the applicability of NIRS models for Snapper <156 months (13 years) and ≥156 months.

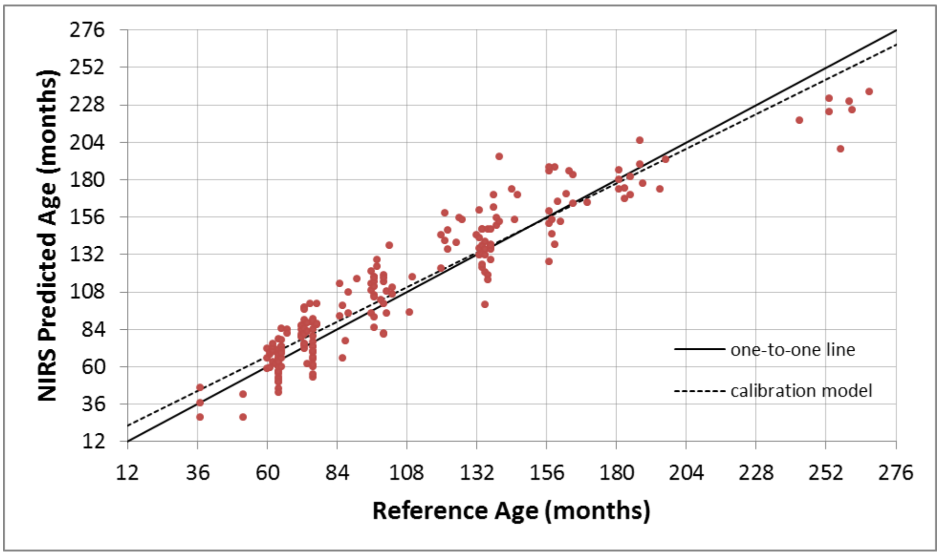


Figure 8. Observed age (reference age) versus NIRS predicted age of Snapper all ages based on a generic wavelength calibration model using otoliths collected in 2012 from the Gulf St Vincent all months combined (n=216, validation set)

The PLS calibration and validation statistics for Snapper otoliths <156 months old and ≥156 months old from the Sunshine Coast and the GSV, based on a generic wavelength selection are presented in Table A3.5 and Table A3.6 respectively. Calibration models developed on otoliths aged <156 months had decreased R^2's but with reductions in RMSE of between three and six months (Table A3.5) compared to Snapper models incorporating all ages (Table A3.4). Reduced age variability in the calibration sample, as indicated by the smaller standard deviation in the age range of the sample, may explain the lower RMSE's. Figure 9 shows the observed versus predicted ages for Snapper <156 months old from the GSV 2012 all months combined.

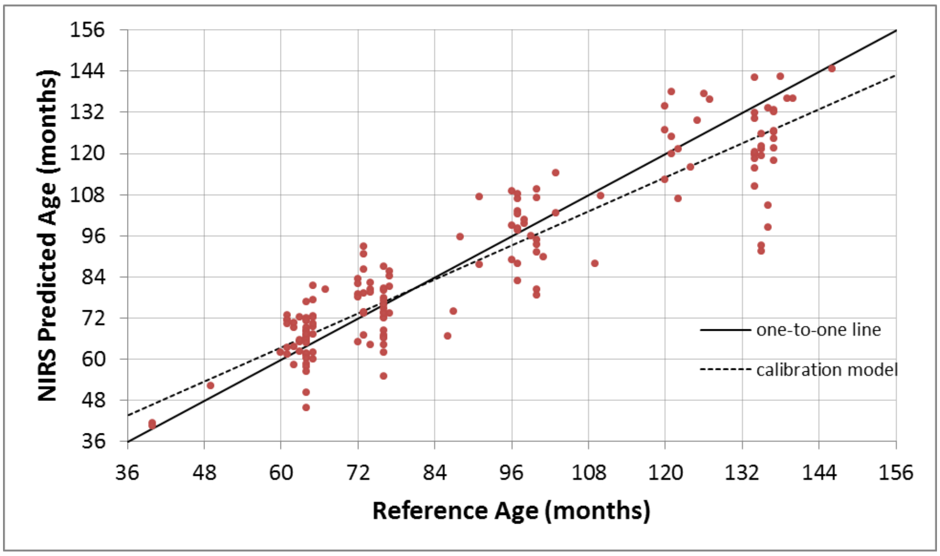


Figure 9. **Observed age (reference age) versus NIRS predicted age of Snapper <156 months based on a generic wavelength calibration model using otoliths collected in 2012 from the Gulf St Vincent all months combined (n=164, validation set)**

Only 26 Snapper otoliths from the Sunshine Coast were ≥156 months old, providing insufficient samples to develop a calibration model for this location.

In total, 76 otoliths were ≥156 months from the GSV pooled across samples from Summer & Autumn 2012 and Winter & Spring 2012, and were evaluated using cross-validation. The cross-validation model had results similar to other Snapper models, with an $R_{cv}^2 = 0.81$, RMSECV = 16.7 months (Table A3.6). Figure 10 shows the observed versus predicted ages of Snapper ≥156 months old from the GSV 2012 all months combined.

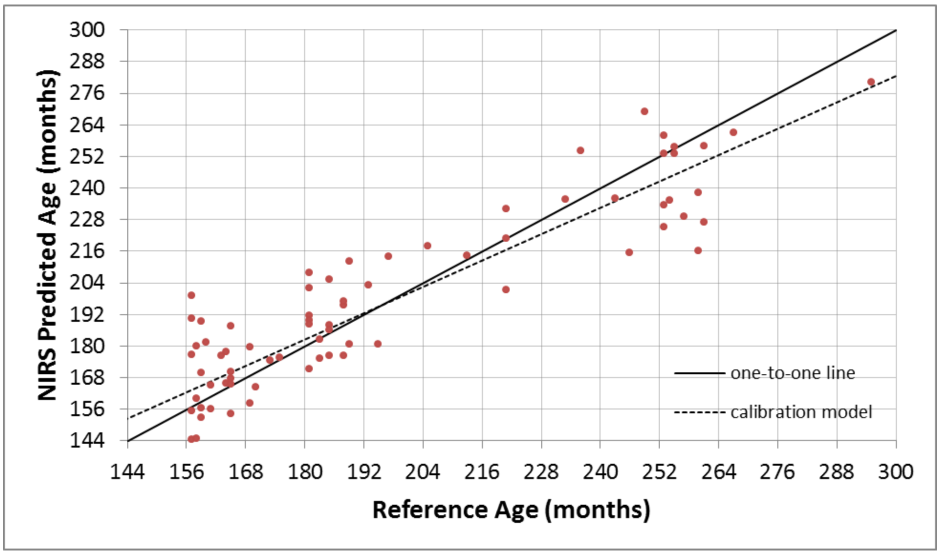


Figure 10. **Observed age (reference age) versus NIRS predicted age of Snapper ≥156 months based on a generic wavelength calibration model using otoliths collected in 2012 from the Gulf St Vincent all months combined (n=76, cross-validation set)**

Geographic variability based on geographic specific wavelength selection

As with the Barramundi models, the predictive performance of NIRS calibration models can be marginally improved for each location by using location-specific NIRS wavelength selection, (Table A3.7). The specific wavelength region relates to spectral differences between locations that probably reflect spatial variability in otolith microchemistry. Subsequently, as this more specific wavelength selection is designed for each location, the model's predictive performance is less than that of the generic wavelength selection model when applied to samples from other geographic locations (data not included).

However, selecting specific wavelengths based on the chemical attributes of scanned otoliths allowed the otoliths collected from different geographic locations to be segregated by Principal Components Analysis (PCA). Figure 11 highlights the separation of GSV Snapper otoliths into two groups with one group exclusively containing otoliths from the Northern GSV when using a geographically specific wavelength selection. This result suggests that NIRS may have applications in the spatial discrimination of fish, such as population/stock structure; marine and freshwater use by life history stages; and links between natal rivers or nursery areas and adult stocks (Thresher 1999). NIRS may be a cost-effective addition to methods such as electron probe microanalysis (EPMA); inductively coupled plasma atomic emission spectrometry (ICP-EAS); proton-induced X- ray emission (PIXE) or Inductively-couple plasma-mass spectrometry (ICP-MS), where a calibration model based on this primary technique could be applied to a much greater number of samples than current costs and research budgets permit.

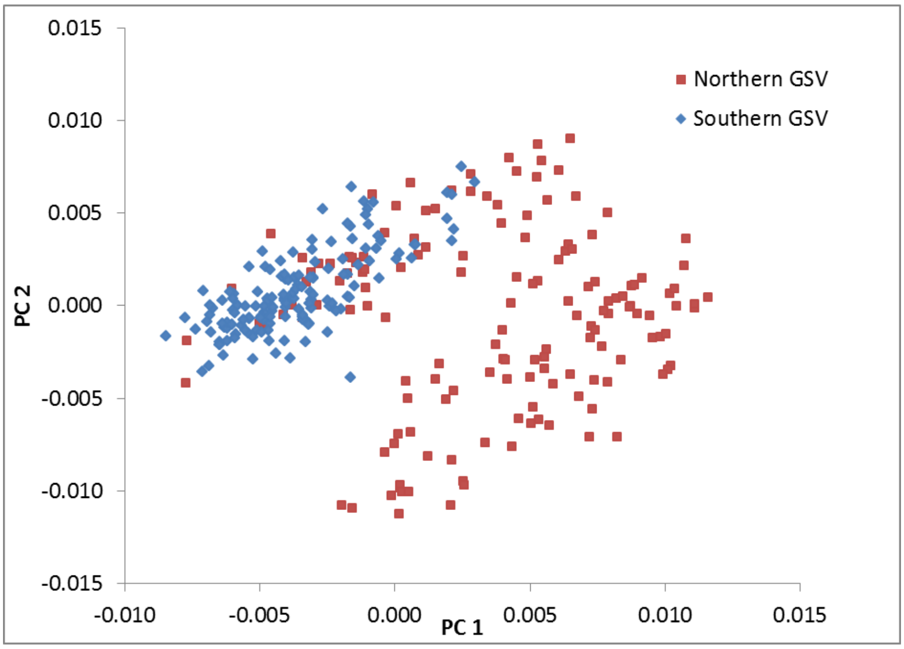


Figure 11. Principal Components Analysis of 2012 Snapper otoliths from the Northern (red squares) and Southern (blue diamonds) Gulf St Vincent, South Australia

Objective 3. Effects of otolith storage time on NIRS

The chemistry of otoliths provides spectral information that is assumed to be related to fish age. Stored otolith samples may suffer slow degradation over time due to oxidation or other environmental effects (e.g., moisture, mildew, light, heat exposure). These factors may result in spectral differences between otoliths over time and between the same otolith scanned at different intervals during storage. If otolith chemistry changes throughout storage, then the NIRS calibration model may be affected. It is possible that the amino acids within otoliths denature over time; pH may change through the loss of fluids and ions as the moisture level stabilises; thus resulting in a spectral change of the otolith from time of collection through subsequent storage.

Short-term storage of otoliths

Short-term storage of otoliths focussed on 'fresh' Barramundi and Snapper otoliths from Queensland that were collected in 2012. NIR spectra were repeatedly collected from these samples up to 17 months after collection (Table 2).

Table 2. NIR spectra acquisition points for fresh otolith samples

Species	Sample Site (Queensland)	Collection Date	NIR Spectra Acquisition Point [a] (Months)
Barramundi	Archer River Estuary	May 2012	3, 4, 6, 9, 12, 15
	Fitzroy River Estuary	February 2012	5, 6, 8, 11, 14, 17
Snapper	Sunshine Coast Offshore	July 2012	3, 6, 12, 18

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

Model performance statistics presented previously in the current report are indicative of a single calibration model constructed from the first NIR spectra acquisition point; i.e., time in months since removal of the otolith from the fish.

Models presented in this section of the report compare the performance of each calibration model for a distinct NIR spectra acquisition point predicting against the remaining NIR spectra acquisition points (i.e., other storage time intervals within each species by location). Calibration and validation data sets were used for the first NIR spectra acquisition point, while full cross-validation was used for the remaining NIR spectra acquisition points. We would expect a lower R^2 and higher RMSE when a calibration model is used to predict a different NIR spectra acquisition point.

Fresh Barramundi otoliths

The R^2 and RMSE model statistics for 'fresh' Barramundi otoliths from the Archer River spanning three to 15 months storage following collection and subsequent NIR spectra acquisition points are presented in Table 3 and Figure 12. The 3-month calibration model was used to predict otolith ages based on the NIR spectra collected at 4, 6, 9, 12 and 15 months following otolith collection. For the 3-month calibration model (i.e., dark blue line, Figure 12), the values plotted at the 3-month NIR spectra acquisition point are the results obtained from the prediction of the 3-month validation set. This procedure was repeated using the 4-month calibration model (i.e., pink line, Figure 12) predicting on NIR spectra acquired at 3, 6, 9, 12 and 15 months and so forth, to indicate predictive model performance in relation to degradation over time. In Figures 12 to 18, the calibration model is represented by the coloured line and the data set predicted is given on the x-axis.

For Archer River Barramundi, the R^2 and RMSE appear to plateau (and remain stable) for NIR spectra acquired from ~6 months following collection for all calibration models except the 3-month model (Figure 12). The R^2 of the 3-month calibration model decreased from 0.88 to 0.79, whilst the RMSE increased from 5.9 to 7.5 months (Table 3). The trend for Archer River calibration statistics to plateau after six months storage (i.e., R^2 from 0.85 to 0.90 and RMSE from 5.1 to 6.3 months) suggests that the otolith chemistry had stabilised after six months.

Table 3. Partial least squares statistics for biological age of 2012 Archer River Barramundi using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 3 to 15 months following otolith collection

R^2 for Each Calibration Model Based on NIR Spectra Acquisition Point [a]							
Calibration Model	**3**	**0.88**	0.85	0.83	0.83	0.79	0.79
	4	0.88	**0.89**	0.89	0.89	0.88	0.88
	6	0.84	0.89	**0.88**	0.88	0.90	0.89
	9	0.88	0.89	0.89	**0.88**	0.89	0.88
	12	0.41	0.75	0.85	0.87	**0.89**	0.89
	15	0.69	0.84	0.88	0.87	0.90	**0.88**
		3	**4**	**6**	**9**	**12**	**15**
		Months After Collection					

RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]							
Calibration Model	**3**	**5.9**	6.4	6.9	6.7	7.4	7.5
	4	5.9	**5.6**	5.4	5.6	5.6	5.6
	6	7.1	5.4	**5.7**	5.6	5.1	5.5
	9	6.3	5.5	5.5	**5.8**	5.5	5.6
	12	13.7	8.1	6.3	5.9	**5.5**	5.4
	15	10.0	6.5	5.7	5.9	5.2	**5.7**
		3	**4**	**6**	**9**	**12**	**15**
		Months After Collection					

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

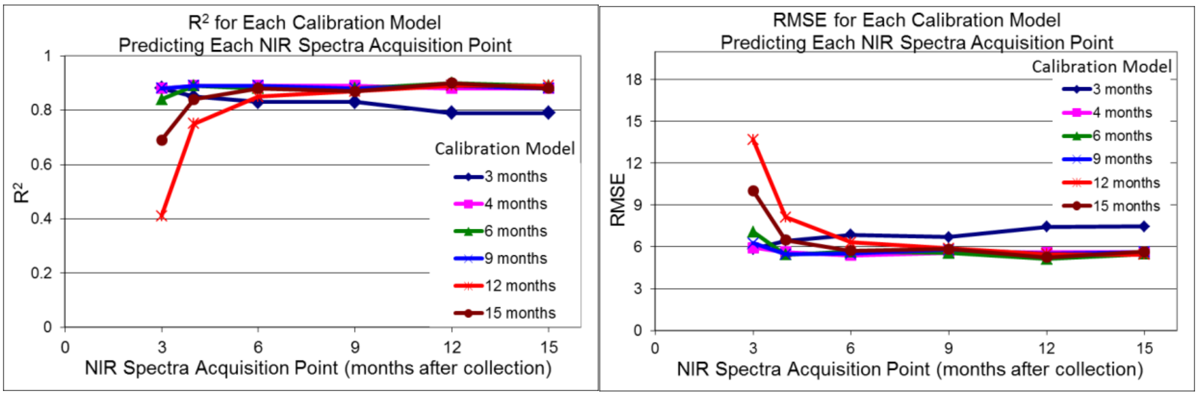


Figure 12. Partial least squares statistics for biological age of 2012 Archer River Barramundi using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 3 to 15 months following otolith collection

In comparison, the fresh Barramundi otolith samples from Fitzroy River stored for between 5 to 17 months following collection and subsequent NIR spectra acquisition appear to stabilise after ~11 months (Table 4 and Figure 13). This can be seen by the 11, 14 and 17 months acquisition point calibration models (i.e., dark blue, red and brown lines in Figure 13) producing reasonably consistent R^2 (0.73 to 0.84) and RMSE's (7.9 to 10.3 months) when predicted against NIRS acquired over the same period (i.e., at 11, 14 and 17 months on the *x*-axis of Figure 13).

The R^2 for the 5-month calibration model when used to predict subsequent NIR spectra acquisition points could not be calculated, indicating a very poor model fit. However, for each of these models the correlation coefficient (r) between the visual assessed age and the NIRS-predicted age was between 0.91 and 0.92, but there were large biases ranging from 21.4 to 41.8 months. The NIRS predictions were improved through bias correction, with R^2 increasing to between 0.78 and 0.83 and RMSE reducing to between 8.2 and 9.3 months. The difference in time to stabilise between otoliths collected from the Archer River and Fitzroy River might be attributed to differences in the collection, processing and storage procedures (Table A2.3) and/or spatial differences influencing the (macro and micro) chemistry of these otoliths.

Table 4. Partial least squares statistics for biological age of 2012 Fitzroy River Barramundi using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 5 to 17 months following otolith collection

R^2 for Each Calibration Model based on NIR Spectra Acquisition Point [a]

		Months After Collection					
		5	**6**	**8**	**11**	**14**	**17**
Calibration Model	**5**	**0.87**	-	-	-	-	-
	6	0.60	**0.82**	0.74	0.67	0.42	0.55
	8	-	0.41	**0.83**	0.66	0.24	0.51
	11	-	-	0.62	**0.82**	0.73	0.83
	14	-	0.30	0.68	0.77	**0.81**	0.81
	17	-	-	0.53	0.84	0.80	**0.82**

RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]

		Months After Collection					
		5	**6**	**8**	**11**	**14**	**17**
Calibration Model	**5**	**7.5**	22.9	32.1	37.2	42.8	38.1
	6	13.6	**8.6**	10.1	11.4	15.0	13.3
	8	46.6	15.2	**8.2**	11.5	17.2	13.8
	11	58.6	24.3	12.1	**8.4**	10.3	8.2
	14	29.9	16.6	11.1	9.5	**8.7**	8.7
	17	58.1	25.6	13.5	7.9	8.8	**8.4**

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

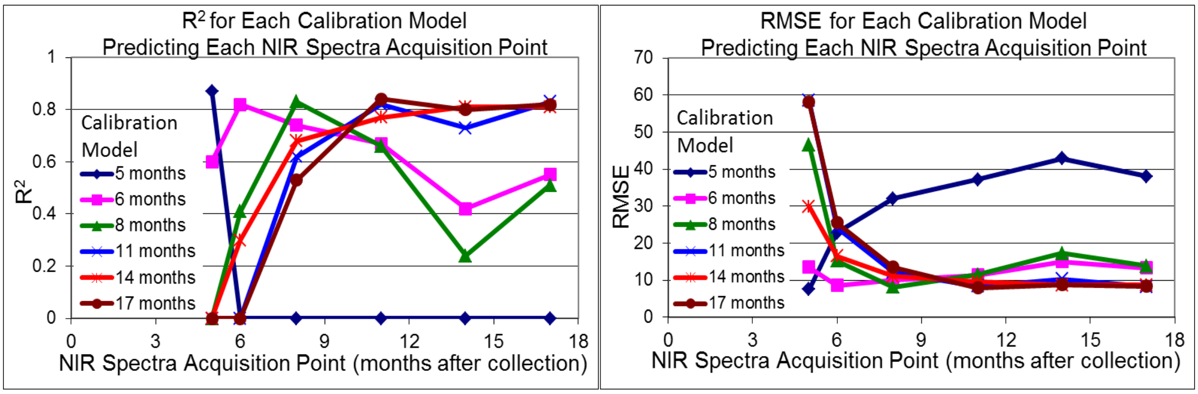


Figure 13. Partial least squares statistics for biological age of Barramundi otoliths collected in 2012 from the Fitzroy River using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 5 to 17 months following otolith collection

Fresh Snapper otoliths

The R^2 and RMSE model statistics for ‘fresh’ Snapper otoliths from the Sunshine Coast spanning 3 to 18 months storage following collection and subsequent NIR spectra acquisition are presented in Table 5 and Figure 14. For Sunshine Coast Snapper, the R^2 and RMSE appear to plateau (and remain stable) for NIR spectra acquired from ~6 months following collection (R^2’s from 0.65 to 0.72 and RMSE from 14.6 to 16.5 months).

The 6-month NIR spectra acquisition point calibration model predicting the NIR data for 18 months after collection, and the 18-month NIR spectra acquisition point calibration model predicting 6 months after collection have similar R^2’s (0.65 and 0.72 respectively) and RMSE values (16.5 and 14.8 months respectively). The stabilisation of calibration statistics after 6 months following otolith collection suggest that otolith chemistry had stabilised, a similar result to that for Archer River Barramundi.

Table 5. Partial least squares statistics for biological age of 2012 Sunshine Coast Snapper using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 3 to 18 months following otolith collection

R^2 for Each Calibration Model Based on NIR Spectra Acquisition Point [a]					
Calibration Model	**3**	**0.69**	0.54	0.53	0.46
	6	0.69	**0.70**	0.71	0.65
	12	0.58	0.71	**0.71**	0.72
	18	0.62	0.72	0.72	**0.70**
		3	**6**	**12**	**18**
		Months After Collection			

RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]					
Calibration Model	**3**	**14.9**	18.7	19.1	20.4
	6	16.6	**15.4**	15.1	16.5
	12	19.2	15.0	**15.2**	14.8
	18	18.3	14.8	14.6	**15.4**
		3	**6**	**12**	**18**
		Months After Collection			

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

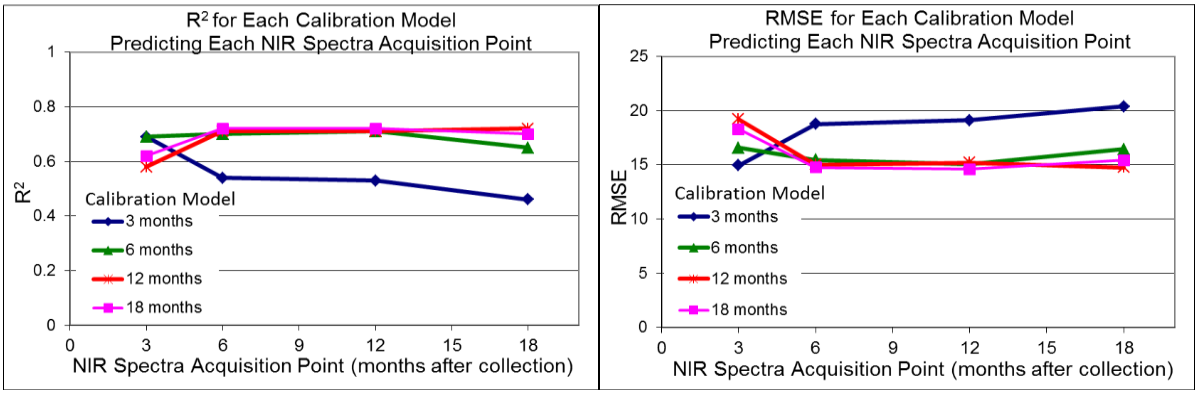


Figure 14. Partial least squares statistics for biological age of 2012 Sunshine Coast Snapper using a generic NIRS wavelength selection, with NIR spectra repeatedly acquired at 3 to 18 months following otolith collection

Long-term storage of otoliths

Long-term storage of otoliths focussed on ‘historic’ Barramundi and Snapper otoliths from Queensland that were collected as far back as 2003 (Table 6). Unlike the short-term storage assessment, we did not compare the repeat scan of the same samples, rather we compared NIRS models based on NIRS scans of otoliths collected from a single location but in different years.

Table 6. NIR spectra acquisition points for historic otolith samples

Species	**Sample Site**	**Collection Date**	**NIR Spectra Acquisition Point [a] (Months)**
Barramundi	Archer River Estuary, Queensland	May 2012	12
		March 2009	48
		April 2006	82
	Fitzroy River Estuary, Queensland	February 2012	11
		February 2009	44
		February 2006	68
		October 2003	105
Snapper	Sunshine Coast Offshore Waters, Queensland	July 2012	12
		July 2009	51
		September 2006	85

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

Historic Barramundi otoliths

The R^2 and RMSE model statistics for the ‘historic’ Barramundi otoliths from the Archer River were highly variable, showing no clear trends indicating further stabilisation or degradation over time (Table 7 and Figure 15). However, the 82-month sample (collected in April 2006) had a large number of otoliths that were discoloured, which may have affected the NIR spectra and subsequently the calibration model. This may explain the low R^2 and higher RMSE when the 12- and 48-month calibration models are used to predict this data set.

Table 7. Partial least squares statistics for biological age of historic Archer River Barramundi using a generic NIRS wavelength selection, with NIR spectra acquired at 12 to 82 months following otolith collection

R^2 for Each Calibration Model Based on NIR Spectra Acquisition Point [a]				
Calibration Model	**12**	**0.89**	0.89	0.29
	48	0.80	**0.94**	-
	82	0.51	0.85	**0.85**
		12	**48**	**82**
		Months After Collection		
RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]				
Calibration Model	**12**	**5.6**	7.3	11.2
	48	7.4	**5.7**	22.5
	82	11.4	8.8	**8.0**
		12	**48**	**82**
		Months After Collection		

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

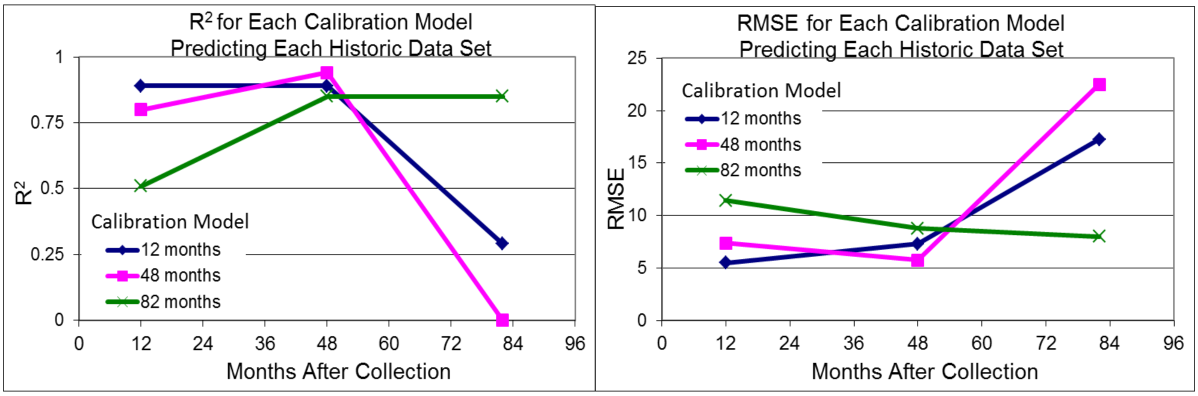


Figure 15. Partial least squares statistics for biological age of historic Archer River Barramundi using a generic NIRS wavelength selection, with NIR spectra acquired at 12 to 82 months following otolith collection

The R^2 and RMSE model statistics for ‘historic’ Barramundi otoliths from the Fitzroy River were also highly variable, with no clear trends to indicate further stabilisation or degradation over time (Table 8 and Figure 16). Some models had low or non-reportable R^2, indicating a very poor model fit. However, for each of these models the correlation between the visual estimated age and the predicted age was between 0.69 and 0.91, but there were large biases ranging from -19.2 to 11.7 months, which could be improved through bias correction. We speculate that the variable results are the consequence of between year differences, as each historic otolith set was obtained from a different year. Temporal effects have major impacts on the robustness of NIRS calibration models and are quite typical of many biological commodities. Temporal variability can be accommodated by incorporating samples from different years into the calibration model.

Table 8. Partial least squares statistics for biological age of historic Fitzroy River Barramundi using a generic NIRS wavelength selection, with NIR spectra acquired at 11 to 105 months following otolith collection

R^2 for Each Calibration Model Based on NIR Spectra Acquisition Point [a]					
Calibration Model	**11**	0.82	-	0.69	0.89
	44	0.73	0.67	0.65	0.85
	68	0.39	-	0.86	0.50
	105	0.79	0.41	-	0.93
		11	**44**	**68**	**105**
		Months After Collection			

RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]					
Calibration Model	**11**	8.4	22.5	9.5	7.3
	44	10.4	7.4	10.0	8.5
	68	15.5	13.0	6.4	15.4
	105	9.1	9.8	14.0	5.9
		11	**44**	**68**	**105**
		Months After Collection			

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

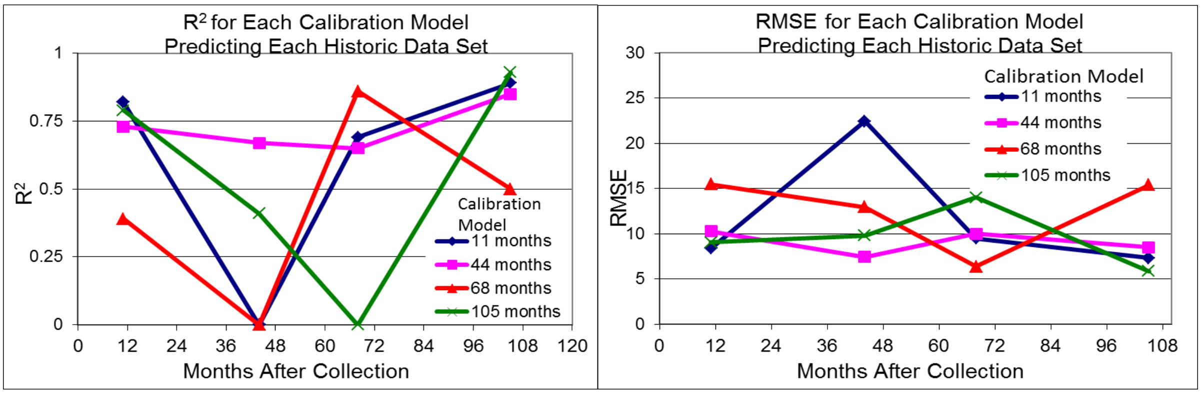


Figure 16. Partial least squares statistics for biological age of historic Fitzroy River Barramundi using a generic NIRS wavelength selection, with NIR spectra acquired at 11 to 105 months following otolith collection

All ‘historic’ Barramundi data sets (i.e., Archer River 12, 48, 82 months and Fitzroy River 11, 44, 68, 105 months) were combined and divided into a calibration and validation set, but selected for fish ≤120 months old. The combined-years Barramundi calibration model (n=328) had an $R_c^2 = 0.82$, RMSEC = 8.4, SDR = 2.34 and bias = 0.001. Prediction statistics for the validation set (n=329) were $R_v^2 = 0.83$, RMSEP = 8.4, SDR = 2.45 and bias = -0.464. Although the calibration models for individual historical data sets did not predict other independent historical data sets well, the results from the combined historical data model suggest that by including both spatial and temporal variability into the model, there is the potential for Barramundi otoliths obtained from different years and locations to be aged using a single calibration model (Figure 17).

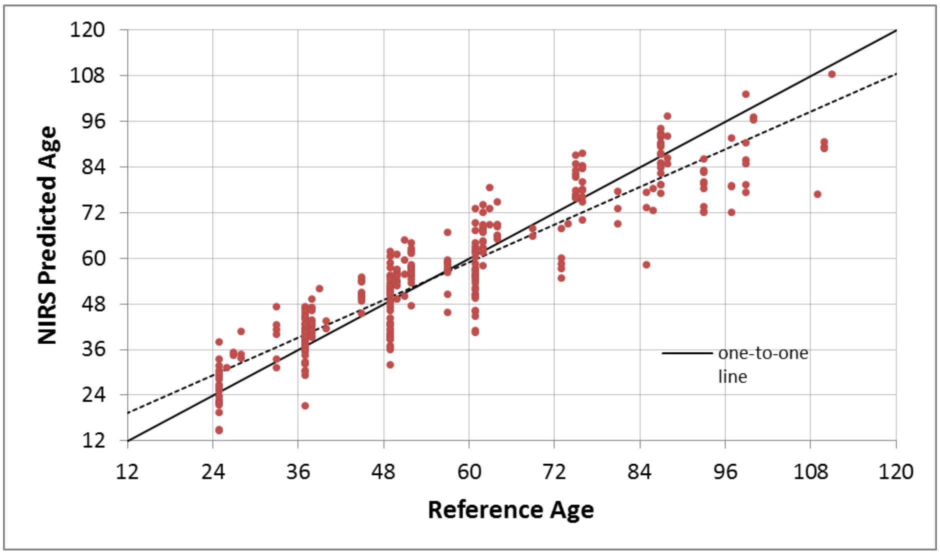


Figure 17. Observed age (reference age) versus NIRS predicted age of Barramundi ≤120 months based on a generic wavelength calibration model using otoliths from Archer and Fitzroy Rivers combined and all collection years combined, (n=329, validation set)

Historic Snapper otoliths

The R^2 and RMSE model statistics for ‘historic’ Snapper otoliths from the Sunshine Coast showed minimal differences between collection years (Table 9 and Figure 18), indicating similar otolith chemistry between samples. One possible explanation for this consistency is the relatively stable environmental conditions of Snapper (i.e., salinity of offshore water) as compared to the variable environment of Barramundi (i.e., salinity of estuarine waters, as well as the use of freshwater habitats by a variable proportion of the Barramundi population).

Table 9. Partial least squares statistics for biological age of historic Sunshine Coast Snapper, using a generic NIRS wavelength selection, with NIR spectra acquired at 12 to 85 months following otolith collection

R^2 for Each Calibration Model Based on NIR Spectra Acquisition Point [a]				
Calibration Model	**12**	0.71	0.80	0.75
	51	0.71	0.82	0.78
	85	0.66	0.72	0.80
		12	**51**	**85**
		Months After Collection		

RMSE (Months) for Each Calibration Model Based on NIR Spectra Acquisition Point [a]				
Calibration Model	**12**	15.2	12.0	14.0
	51	15.0	11.3	13.1
	85	16.2	14.0	12.6
		12	**51**	**85**
		Months After Collection		

[a] Acquisition Point refers to the time since the otolith was collected/removed from the head of the fish to when the NIR spectra was acquired.

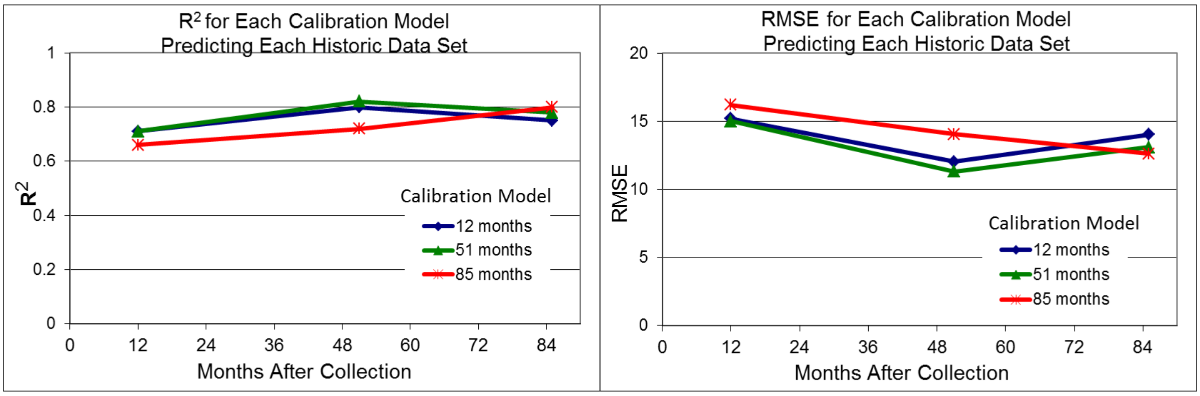


Figure 18. Partial least squares statistics for biological age of historic Sunshine Coast Snapper using a generic NIRS wavelength selection, with NIR spectra acquired at 12 to 85 months following otolith collection

Efficacy of NIRS to predict fish age from a fisheries perspective

The efficacy of ageing fish based on NIRS calibration/prediction models was assessed by comparing the following fisheries performance measures between observed and predicted values for models for Barramundi ≤120 months old and Snapper <156 months old:

(i) the percentage of a sample set that was correctly allocated to its observed age-class/group;
(ii) the percentage of a sample whose predicted age was ±6 months of its observed age;
(iii) the percentage of a sample whose predicted age was ±12 months of its observed age;
(iv) the visual trends in age-bias plots of predicted values;
(v) the Index of Average Percent Error (IAPE), a standard fish ageing metric of ‘precision; and
(vi) statistical differences in age-class distribution using a Kolmogorov-Smirnov test.

Barramundi

Archer River

Archer River Barramundi displayed large variability in length-at-age (Figure 19) that is typical for Barramundi (Halliday and Robins 2007; Halliday *et al.* 2014).

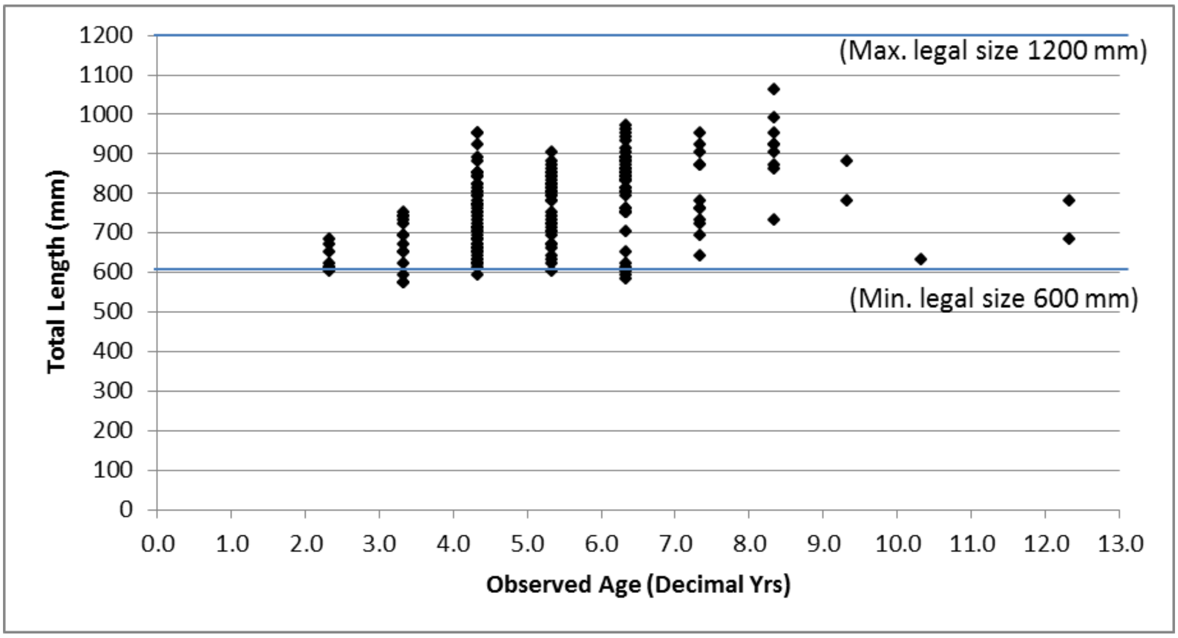


Figure 19. Length-at-age plot of 2012 Archer River Barramundi otolith samples

All NIRS models predicting Barramundi age had high R^2 (>0.88) and relatively small RMSE (≤6.3 months). The R^2 quoted for the NIRS models is the percentage of the total variance accounted for by the explained variance for the given number of latent variables in the PLS regression model, while the R^2 for the otolith weight (and total fish length) models is the adjusted R^2 from regression model. Therefore, the R^2’s quoted in the following table are not directly comparable between the NIRS models and the regression models. Age-class derived from NIRS predicted age was correct for 73% of the 2012 Archer River validation set (Table 10). For the 2012 Archer River validation set, 73% of individuals were predicted to within 6 months of their observed age; 94% of individuals were predicted to within 12 months of their observed age and the IAPE (i.e., between method precision) was 3.4%. Results were similar for the cross-validated short-term storage models; except for the 9-month NIRS acquisition point model which had a lower percentage of correct age-class classification

(Table 10) and a higher IAPE (i.e., 5.2%). Linear regressions (LR) of otolith weight (OtoWt) or multiple linear regression (MLR) of otolith weight and total fish length (TL) had R^2's ranging from 0.43 to 0.57, higher RMSE's, lower percentages of individuals correctly assigned to their observed age-class, or age estimated to within 6 or 12 months of their observed age than models based on NIRS, as well as higher IAPE's (Table 10).

Table 10. Performance of models predicting fish age using NIRS and regression analysis for 2012 Archer River Barramundi

Data Set	Model	*R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Calibration N=92 Ages = 28 - 124 months SD = 18.8 months	NIRS 3 month AP, LV=3	0.89	6.3	-0.00	0.94	0.89	68	73	91	3.7
	LR OtoWt	0.43	14.0				43	49	63	
	MLR OtoWt +TL	0.56	12.3				46	42	70	7.6
2012 Validation N=90 Ages = 28 - 112 months SD = 17.1 months	NIRS 3 month AP, LV=3	0.88	5.9	1.42	0.95	0.96	73	73	94	3.4
	LR OtoWt	0.53	11.8				47	47	72	
	MLR OtoWt +TL	0.57	11.3				44	46	72	7.1
2012 Cross-Validation N=98 Ages = 28 - 112 months SD = 16.5 months	NIRS 4 month AP, LV=3	0.89	5.6	-0.05	0.94	0.89	72	77	96	3.6
	NIRS 6 month AP, LV=3	0.88	5.6	-0.06	0.94	0.89	74	76	94	3.5
	NIRS 9 month AP, LV=3	0.88	5.8	-0.05	0.94	0.88	59	58	89	5.2
	NIRS 12 month AP, LV=3	0.89	5.5	-0.06	0.94	0.89	77	76	97	3.3
	NIRS 15 month AP, LV=3	0.88	5.7	-0.03	0.94	0.89	77	76	94	3.5

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables; LR = Linear Regression; MLR = Multiple Linear Regression

The R^2 for NIRS models is the percentage of the total variance accounted for by the explained variance for the given number of latent variables in the model. It is not the same as the square of the correlation coefficient. It is possible for NIRS models to have a low R^2 but a high correlation coefficient (r), suggesting a strong linear relationship is present but the predicted values are not close to the 1:1 equivalence line. This can occur when an NIRS model has large bias across the full age range. The R^2 for the otolith weight models is the R^2 from the generalised linear model adjusted for the number of variable fitted in the model. The R^2 for the NIRS model is not directly comparable to the adjusted R^2 for the otolith weight models.

For any observed age, NIRS and MLR predicted ages were variable (Figure 20), but there was no pattern of age-bias in NIRS predicted age for Archer River Barramundi (Figure 21). Age-bias plots indicated that age predicted from a MLR of otolith weight and total fish length was biased, with age overestimated in younger fish and underestimated in older fish (Figure 21).

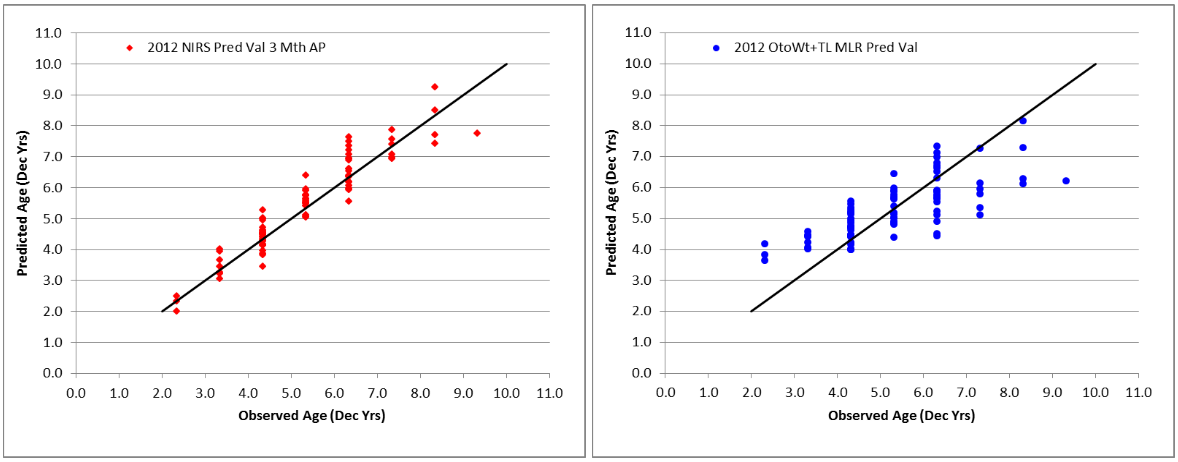


Figure 20. Observed age versus predicted age for 2012 Archer River Barramundi validation set, solid black line indicates the 1:1 equivalence

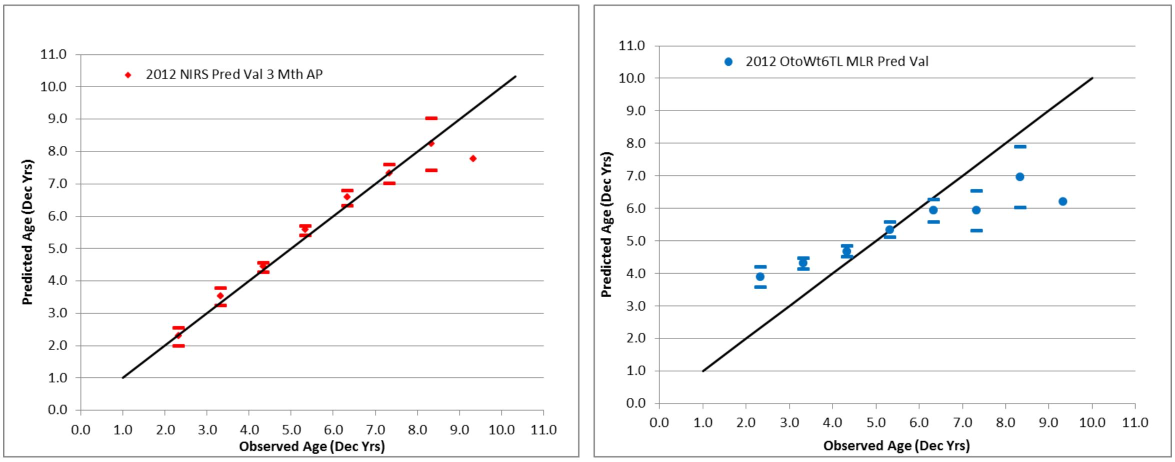


Figure 21. Age-bias plots for 2012 Archer River Barramundi validation set; error markers represent the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence

Age-class distributions for 2012 Archer River Barramundi were similar (Figure 22) and not statistically different ($\chi^2_{(2\text{ d.f.})}$ = 0.200; *p*=0.905) for observed age and NIRS predicted age based on the 3-month acquisition point. However, the age-class distribution based on age predicted from otolith weight and total fish length was significantly different from the observed age-class distribution ($\chi^2_{(2\text{ d.f.})}$ = 6.422; *p*=0.040). The weight of evidence across the fisheries performance measures indicated that NIRS estimates of fish age based on calibration models of otoliths ≤120 months offered good potential to supplement standard fish ageing methods for Archer River Barramundi.

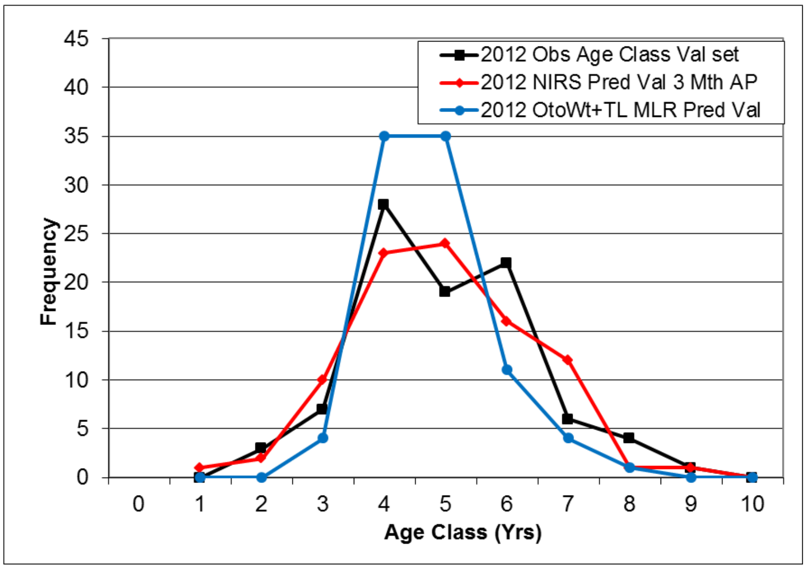


Figure 22. Observed and predicted age-class distributions for 2012 Archer River Barramundi validation set

Two 'historic' sets of Barramundi otoliths from the Archer River collected in 2009 and 2006 were assessed for NIRS ageing of fish. Like the 2012 data, Barramundi in these data sets showed highly variable length-at-age (Figure 23).

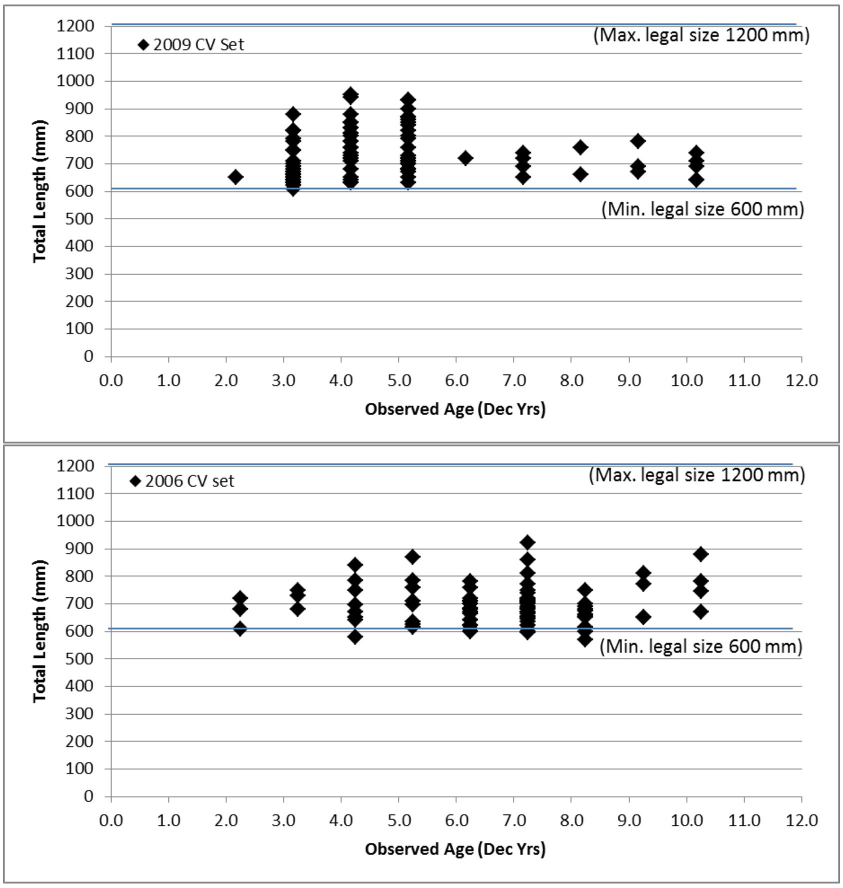


Figure 23. Length-at-age plot of historic Archer River Barramundi otolith samples collected in: (a) March 2009, n=99; and (b) April 2006, n=95

These historic sets of Archer River Barramundi otoliths were assessed for storage degradation and performance using full cross-validation (CV). All models predicting Barramundi age based on NIRS had high R^2 (>0.85) and relatively small RMSE (≤8 months). The NIRS-based age-class was correct for 77% of the 2012 Archer River CV set, with 76% of individuals predicted to within 6 months of their observed age, 97% of individuals predicted to within 12 months of their observed age and an IAPE of 3.3% (Table 11). Samples from 2009 and 2006 had lower percentages of individuals whose age-class was correctly derived (Table 11).

Predicted age using a MLR of otolith weight and total fish length had R^2’s ranging from 0.58 to 0.68, higher RMSE’s, lower percentages of individuals correctly assigned to age-class, lower percentages of individuals predicted within 6 months and 12 months of their observed age and higher IAPE’s than models based on NIRS (Table 11).

Table 11. Performance of models predicting fish age using NIRS and regression analysis for historic Archer River Barramundi

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Cross-Validation N=98	NIRS 12 month AP, LV=3	0.89	5.5	-0.06	0.94	0.89	77	76	97	3.3
Ages = 28 – 112 months SD = 16.5 months	MLR OtoWt +TL	0.59	10.5				48	47	72	6.8
2009 Cross-Validation N=99	NIRS 48 month AP, LV=3	0.94	5.7	-0.02	0.97	0.93	60	72	96	3.9
Ages = 26 – 122 months SD = 22.5 months	MLR OtoWt +TL	0.68	12.6				31	31	64	8.9
2006 Cross-Validation N=95	NIRS 82 month AP, LV=3	0.85	8.0	-0.02	0.92	0.86	54	58	88	4.1
Ages = 27 – 123 months SD = 20.6 months	MLR OtoWt +TL	0.58	13.1				33	36	65	7.3

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables; MLR = Multiple Linear Regression

For any observed age, NIRS and MLR predicted ages were variable (Figure 24).

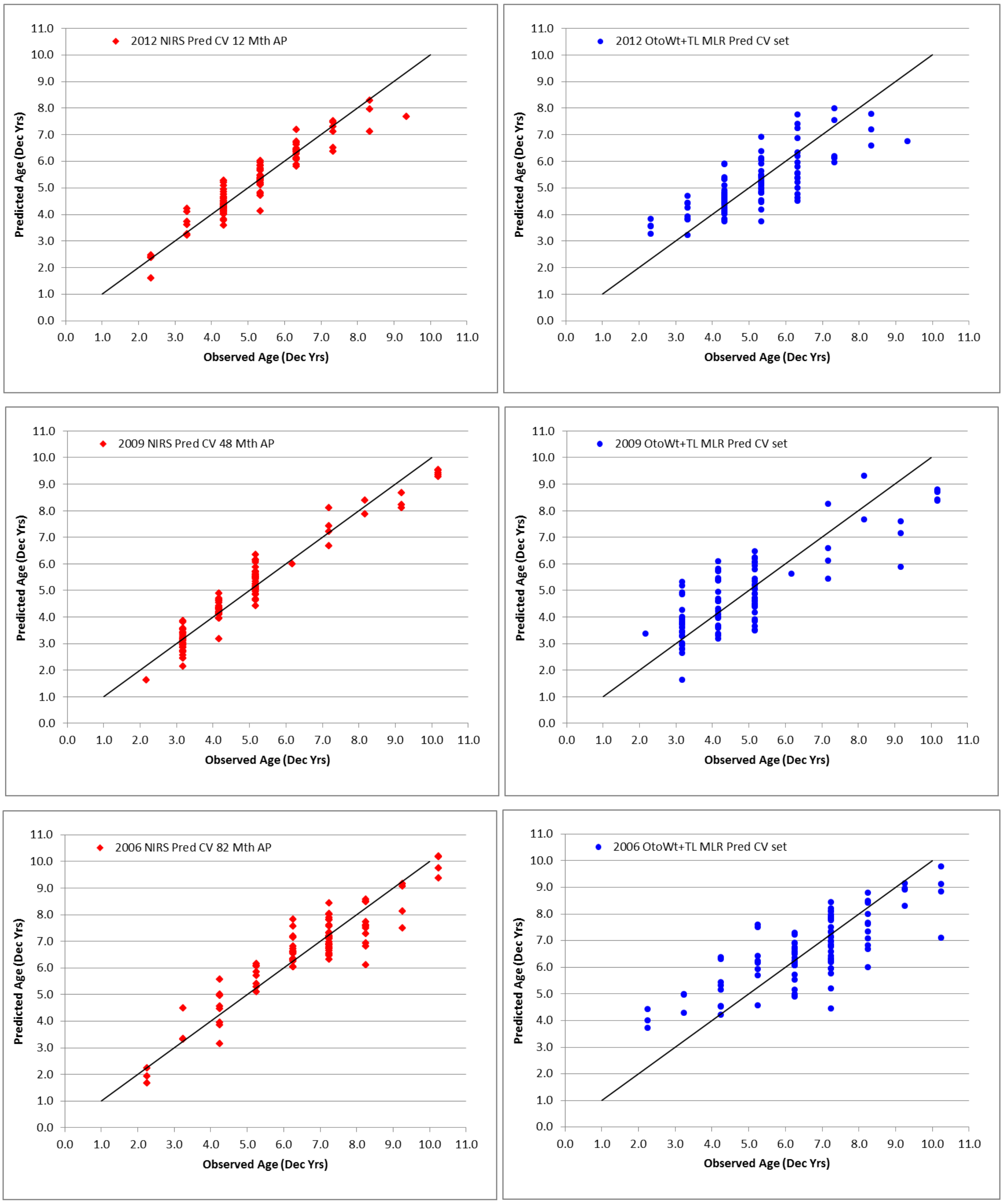


Figure 24. Observed age versus predicted age for historic Archer River Barramundi cross-validation sets; solid black line indicates the 1:1 equivalence

There was no consistent pattern of bias in the NIRS predicted age for Archer River Barramundi, but there was bias in fish ages predicted from a MLR of otolith weight and total fish length (Figure 25).

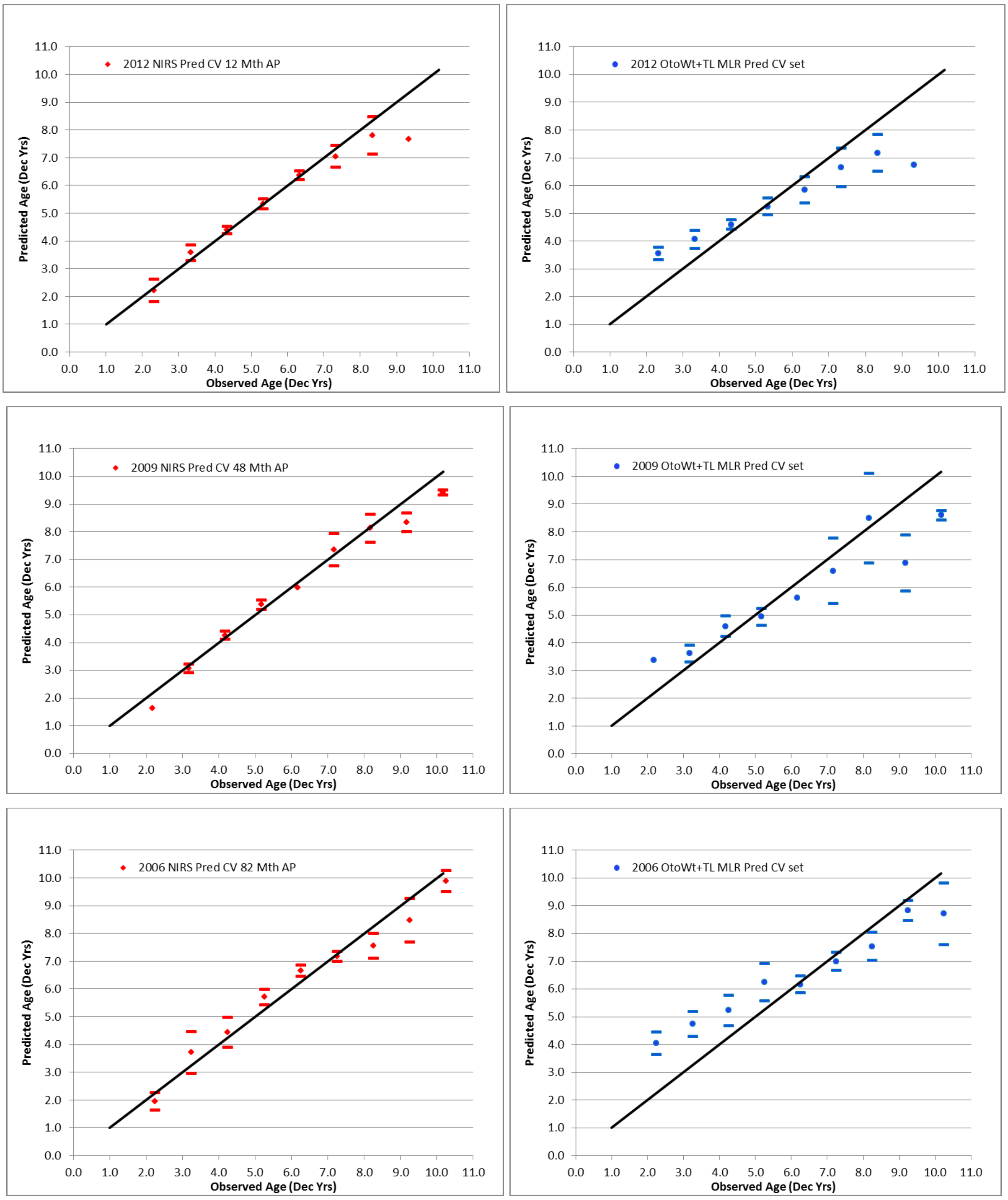


Figure 25. Age-bias plots for historic Archer River Barramundi cross-validation sets; error markers represent the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence

The similarity of the age-class distributions for the cross-validated sets varied between collection years (Figure 26). Despite some visual differences between the observed age-class distribution and the predicted age-class distributions in some years, the differences between observed and NIRS predicted, the observed and MLR predicted, and the NIRS predicted and MLR predicted were not statistically significant i.e., recorded *p*-values were >0.05.

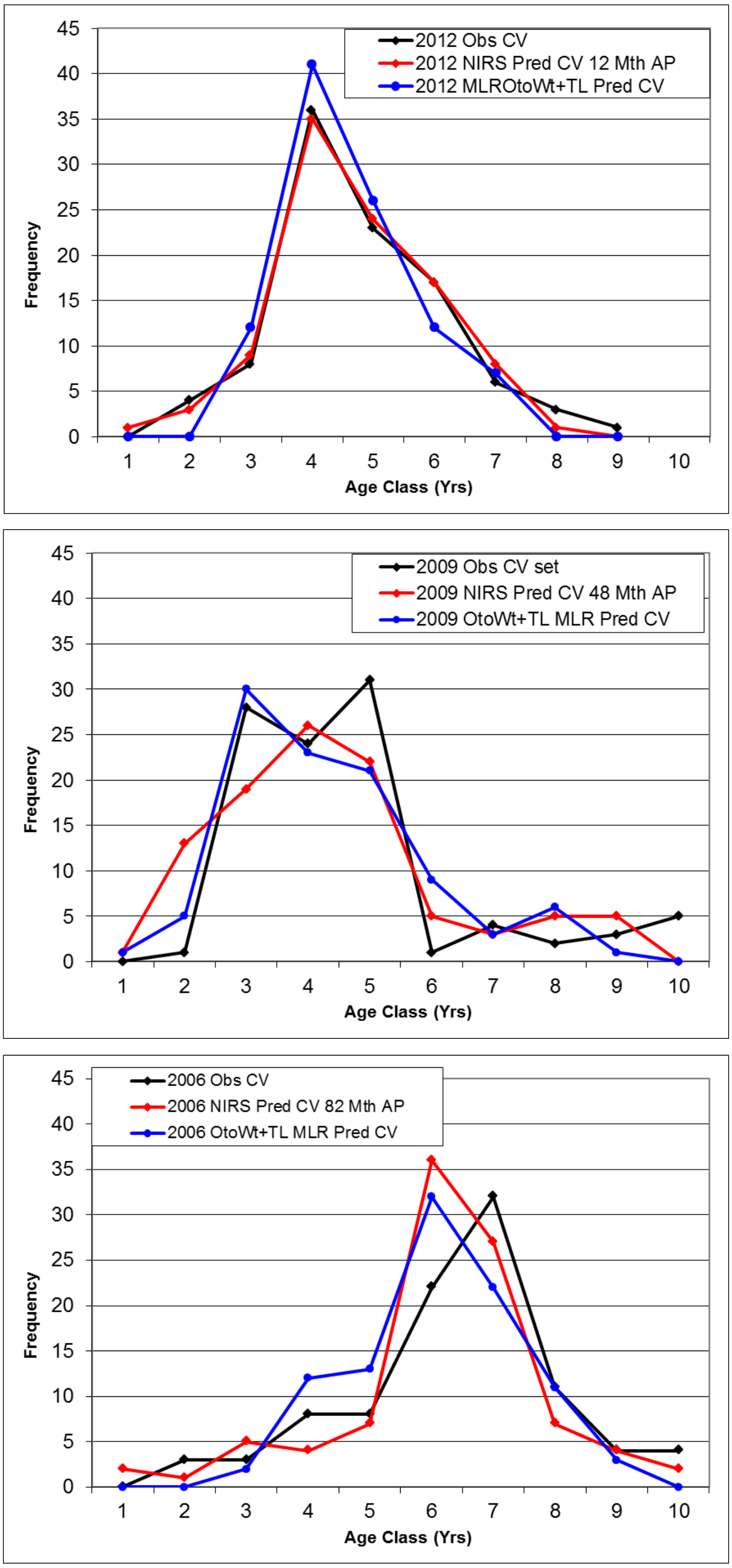


Figure 26. **Observed and predicted age-class distributions for historic Archer River Barramundi cross-validation sets**

Results of fisheries performance measures from the analysis of historic Archer River samples support the potential for NIRS to supplement standard fish ageing methods for Archer River Barramundi.

Fitzroy River

The 2012 samples of Barramundi collected from the Fitzroy River displayed typical variability in length-at-age (Figure 27), with few six and seven year old fish in the sample as a consequence of weak year classes.

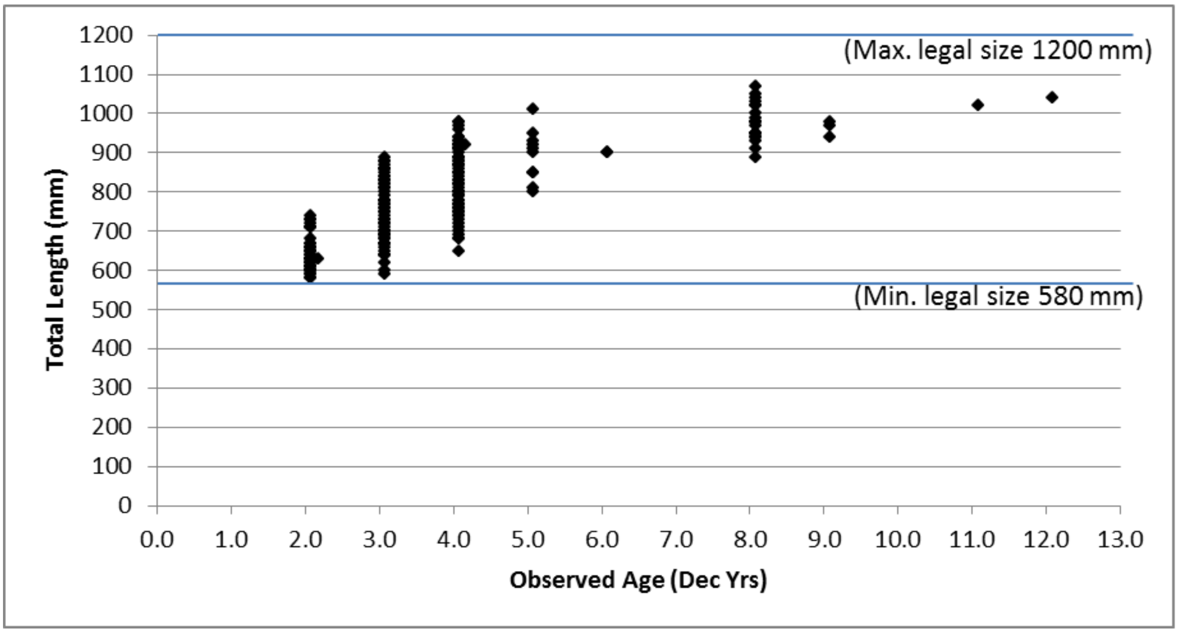


Figure 27. Length-at-age plot of 2012 Fitzroy River Barramundi otolith samples

All models predicting Barramundi age based on NIRS had R^2's greater than 0.80 and RMSE's ≤8.7 months (Table 12). Age-class derived from NIRS predicted age was correct for ~40% of the 2012 Fitzroy River validation set based on a calibration model for the 5-month acquisition point. The 5-month acquisition point corresponds to the first NIR spectra collection (i.e. 'fresh' samples). For this validation set, 57% of individuals were predicted to within 6 months of their observed age; 91% of individuals were predicted to within 12 months of their observed age and the IAPE was 6.9%. Results for all other cross-validation models were similar, but with ~50% of individuals allocated to the correct age-class based on NIRS predicted age and IAPE's ≥7.3%. Regressions of otolith weight or otolith weight and total fish length had R^2's ranging from 0.72 to 0.75, larger RMSE's, lower percentages of individuals correctly assigned to their observed age-class, or age estimated to within 6 or 12 months of their observed age and higher IAPE's than models based on NIRS (Table 12).

Table 12. Performance of models predicting fish age using NIRS and regression analysis for 2012 Fitzroy River Barramundi

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Calibration N=102 Ages = 25 – 109 months SD = 22.2 months	NIRS 5 month AP, LV=4	0.86	8.3	-0.18	0.93	0.88	55	65	87	6.9
	LR OtoWt (3 OR)	0.70	12.1				35	34	71	10.0
	MLR OtoWt +TL	0.70	12.0				33	38	69	10.0
	Power Reg OtoWt (3 OR)	0.73	11.3				39	23	41	19.9
2012 Validation N=100 Ages = 25 – 109 months SD = 20.9 months	NIRS 5 month AP, LV=4	0.87	7.5	-0.24	0.94	0.94	39	57	91	6.9
	LR OtoWt	0.72	11.0				37	41	78	9.2
	MLR OtoWt +TL	0.72	11.1				34	35	74	10.0
	Power Reg OtoWt	0.75	10.3				42	27	46	21.1
2012 Cross Validation N=97 Ages = 25 – 109 months SD = 19.8 months	NIRS 6 month AP, LV=3	0.82	8.6	-0.00	0.90	0.82	51	55	88	8.6
	NIRS 8 month AP, LV=4	0.83	8.2	-0.11	0.91	0.84	49	53	89	7.5
	NIRS 11 month AP, LV=4	0.82	8.4	-0.10	0.91	0.84	51	59	88	7.3
	NIRS 14 month AP, LV=3	0.81	8.7	0.04	0.90	0.82	48	50	89	7.4
	NIRS 17 month AP, LV=4	0.82	8.4	-0.09	0.90	0.84	49	57	90	7.4

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables; OR = Outliers Removed; LR = Linear Regression; MLR = Multiple Linear Regression; Power Reg.: age = $a+b*OtoWt^c$

For any observed age, NIRS and LR predicted ages were variable (Figure 28).

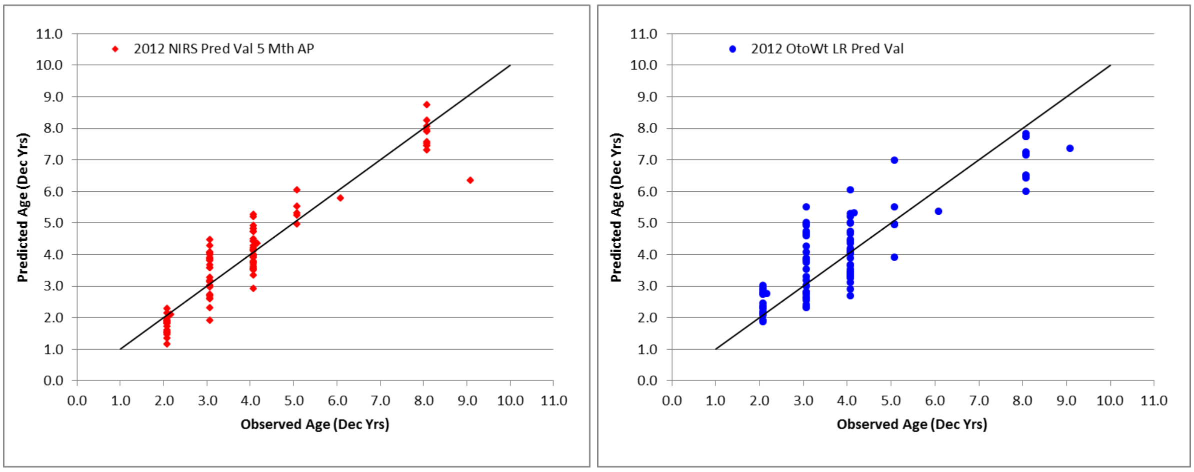


Figure 28. Observed age versus predicted age for 2012 Fitzroy River Barramundi validation set; solid black line indicates the 1:1 equivalence

Age-bias plots indicated no consistent pattern of bias in the NIRS predicted age for Fitzroy River Barramundi (Figure 29). There was some level of bias in age predicted from a linear regression of otolith weight, with a slight pattern of over-estimating age in younger fish and under-estimating age in older fish (Figure 29).

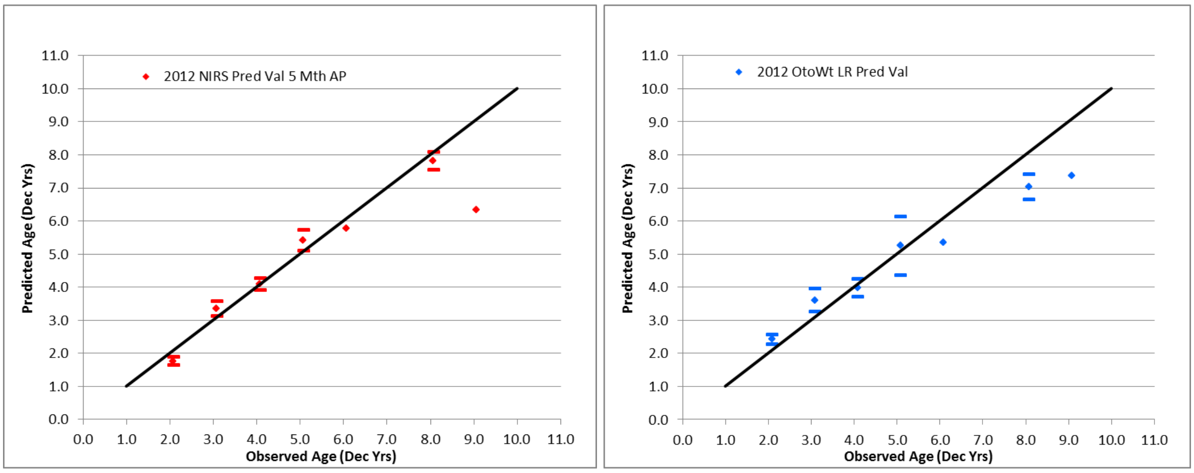


Figure 29. Age-bias plots for 2012 Fitzroy River Barramundi validation set; error markers represent the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence

Age-class distributions were statistically different between the observed age and NIRS predicted age for the 5-month acquisition point ($\chi^2_{(2\text{ d.f.})} = 6.48$; p=0.039); but not between the observed age and the otolith weight predicted age-class distributions ($\chi^2_{(2\text{ d.f.})} = 2.88$; p=0.237); nor between the NIRS predicted and otolith weight predicted age-class distributions ($\chi^2_{(2\text{ d.f.})} = 5.12$; p=0.077) (Figure 30).

Overall, the fisheries performance measures indicated that NIRS estimates of fish age based on calibration models of otoliths ≤120 months had reasonable efficacy, but performance was not as good as for Archer River Barramundi.

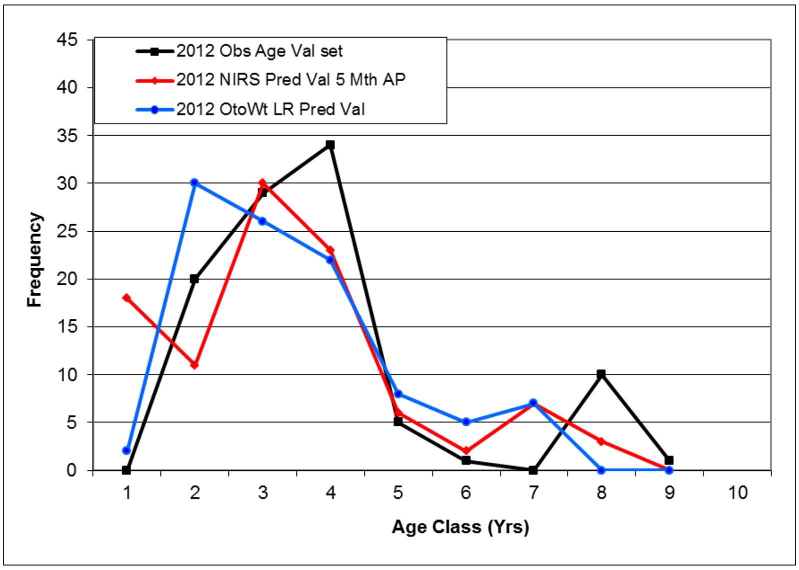


Figure 30. Observed and predicted age-class distributions for 2012 Fitzroy River Barramundi validation set

Three historic sets of Barramundi otoliths from the Fitzroy River estuary were assessed for storage degradation and performance using cross-validation (CV). Otoliths were collected in February 2009 (n=95), February 2006 (n=97) and October 2003 (n=81). Like the 2012 data, Barramundi in these data sets showed highly variable length-at-age (Figure 31).

Models predicting Barramundi age based on NIRS had R^2's ranging from 0.67 to 0.93 and RMSE's ≤8.4 months (Table 13). Age-class derived from NIRS predicted age was correct for 51% of the 2012 Fitzroy River CV set, with 59% of individuals predicted to within 6 months of their observed age, 88% of individuals predicted to within 12 months of their observed age. IAPE was 7.3%. Samples from 2009 and 2006 had similar fisheries efficacy, whilst the 2003 sample had slightly better efficacy (Table 13). Age was predicted using a linear regression of otolith weight (OtoWt) against fish age because for the Fitzroy samples, the additional term of total length was either not statistically significant in the regression model or only marginally improved model fit (e.g., adjusted R^2). Across all CV sets for historic sample years, age predicted from linear regression of otolith weight had lower R^2's, higher RMSE's, lower percentages of individuals correctly assigned to age-class, or within 6 or 12 months of their observed age and higher IAPE's than models based on NIRS (Table 13).

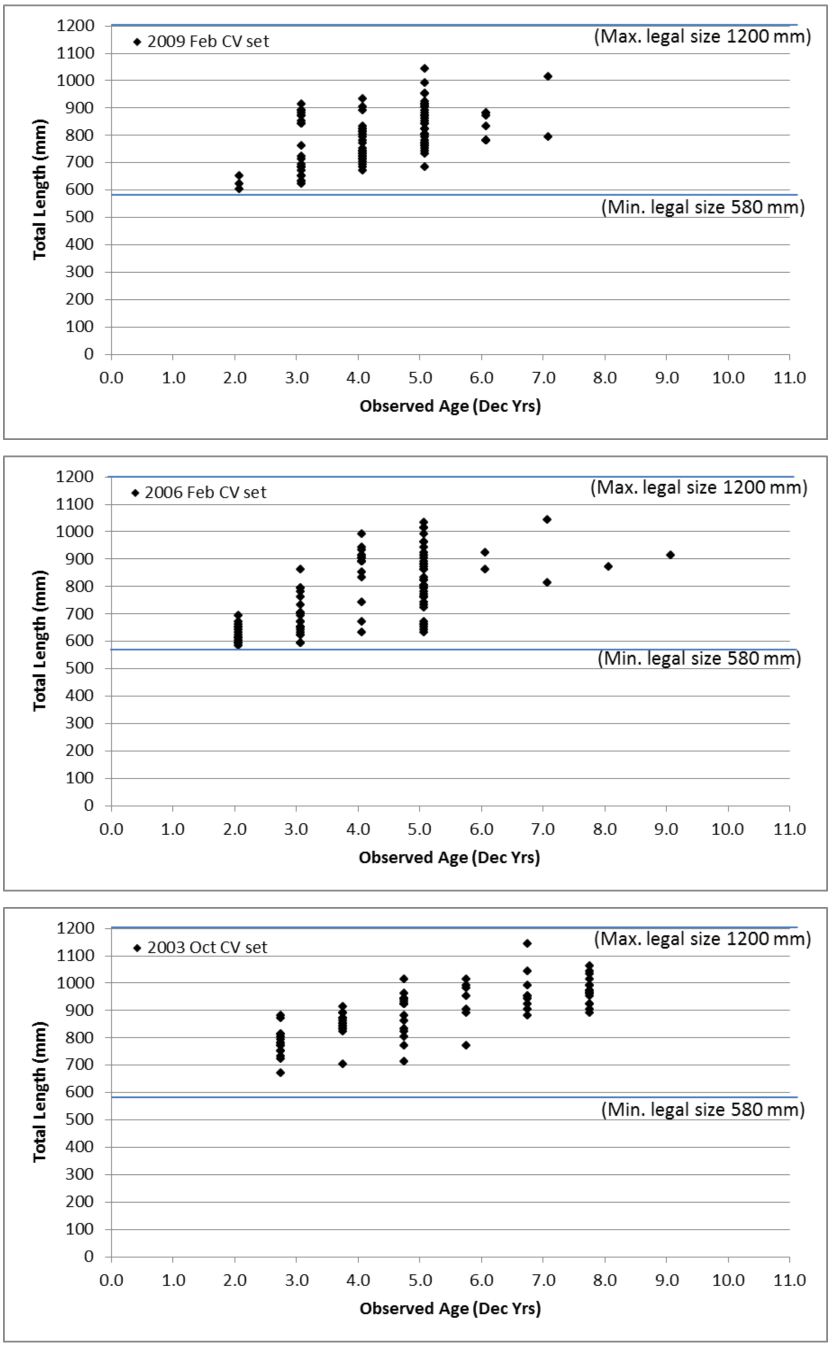


Figure 31. Length-at-age plots of historic Fitzroy River Barramundi otoliths samples collected in: (a) February 2009, n=95; (b) February 2006, n=97; and (c) October 2003, n=81

Table 13. Performance of models predicting fish age using NIRS and regression analysis for historic Fitzroy River Barramundi

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Cross Validation N=97	NIRS 11 month AP, LV=4	0.82	8.4	-0.10	0.91	0.84	51	59	88	7.3
Ages = 25 – 109 months SD = 19.8 months	LR OtoWt	0.60	12.5				37	28	73	10.7
2009 Cross Validation N=95	NIRS 44 month AP, LV=3	0.67	7.4	-0.04	0.81	0.68	47	59	91	5.6
Ages = 25 – 85 months SD = 1.85 months	LR OtoWt	0.33	10.5				33	49	74	8.3
2006 Cross Validation N=97	NIRS 68 month AP, LV=3	0.86	6.4	-0.18	0.92	0.86	51	75	92	5.1
Ages = 25 – 109 months SD = 20.6 months	LR OtoWt	0.49	11.9				42	35	72	9.9
2003 Cross Validation N=81	NIRS 105 month AP, LV=3	0.93	5.9	0.00	0.96	0.93	63	69	96	4.1
Ages = 33 – 93 months SD = 21.9 months	LR OtoWt	0.54	14.6				18	26	44	10.4

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables; OR = Outliers Removed; LR = Linear Regression

For any observed age, predicted age was variable (Figure 32).

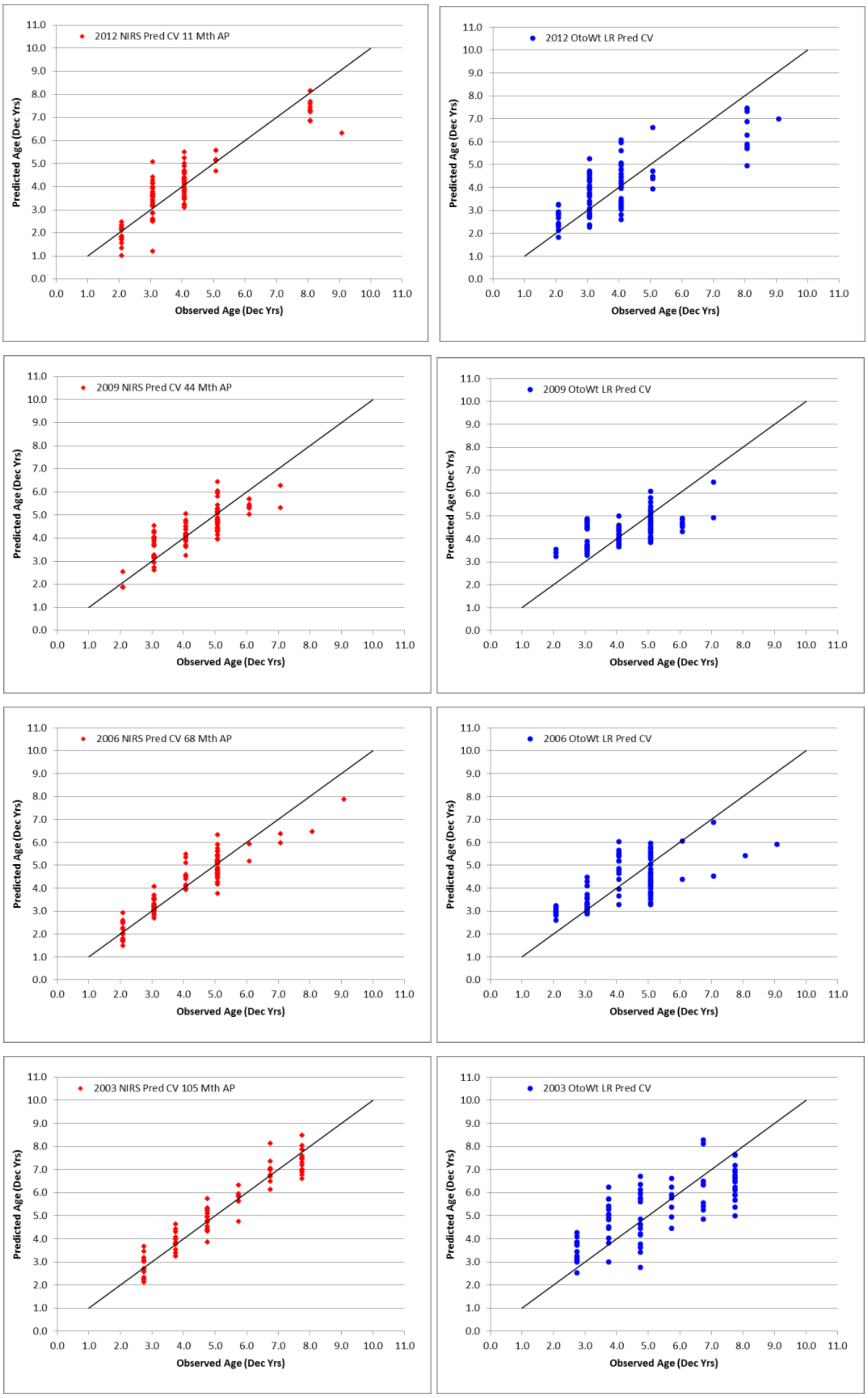


Figure 32. Observed age versus predicted age for historic Fitzroy River Barramundi cross-validation sets, solid black line indicates the 1:1 equivalence

Age-bias plots indicated that, despite models being restricted to fish ≤120 months, age predicted from NIRS had a small amount of bias, underestimating the age of older fish in three of the four historic data sets (Figure 33). However, age predicted from the linear regression of otolith weight had much greater levels of bias, with a consistent pattern of over estimating age in younger fish and underestimating age in older fish (Figure 33). The age-bias plot with the best fit was 2003 NIRS. It maybe noteworthy that this data set had a slightly smaller age range (Table 13) but importantly a more uniform distribution of samples across the age range (i.e., sample numbers >10 in the age groups at the either end of the age range modelled).

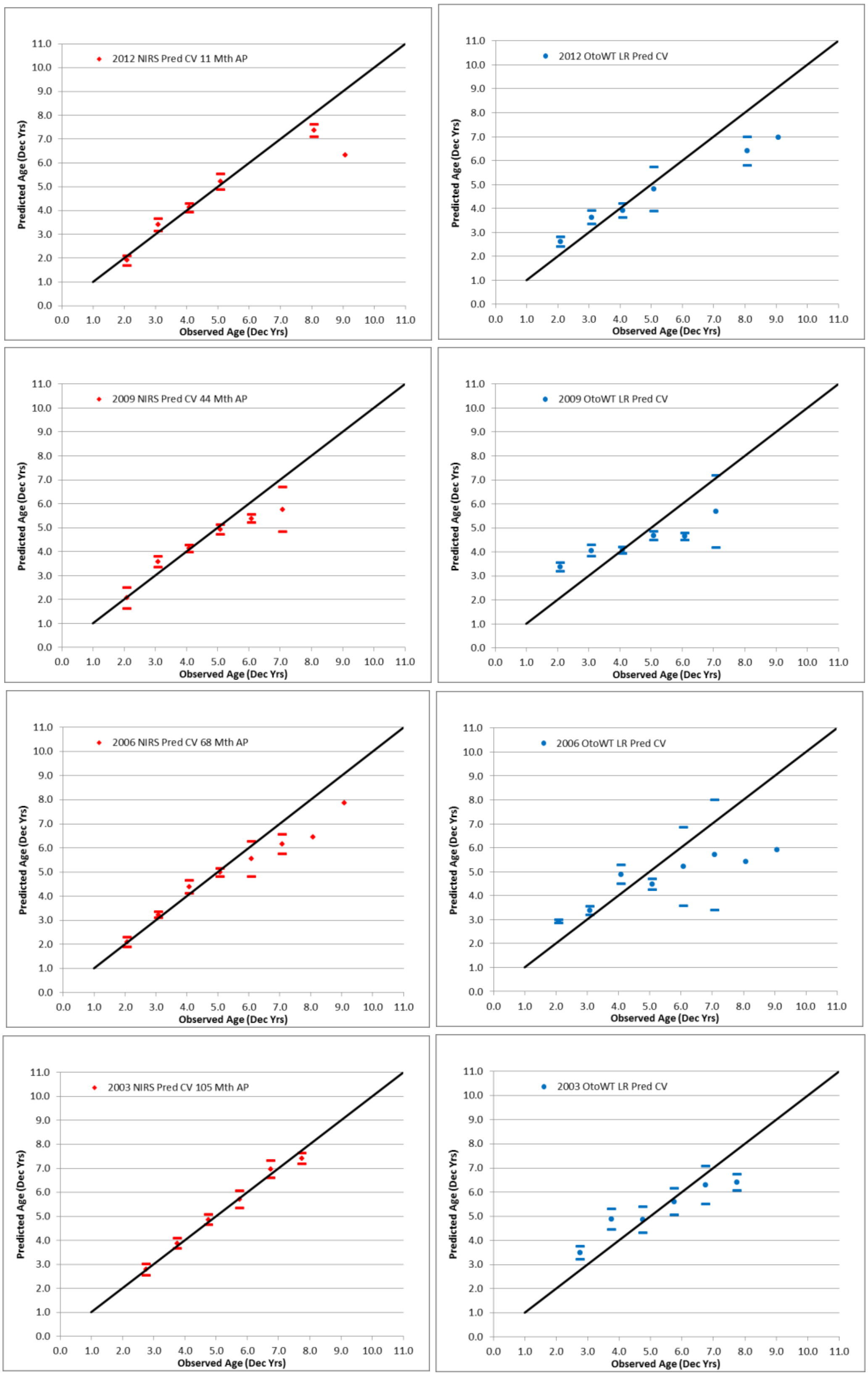


Figure 33. Age-bias plots for historic Fitzroy River Barramundi cross-validation sets; error markers represent the 95% CI around the predicted mean of each observed age group; solid black line indicates the 1:1 equivalence

The similarity of the age-class distributions for the CV sets varied between collection years (Figure 34). In only three comparisons, were the distributions significantly different. These were: (i) the age-class distribution predicted from otolith weight against the observed age-class for 2009 ($\chi^2_{(2\ \mathrm{d.f.})} = 22.93$; $p < 0.001$); (ii) the age-class distribution predicted from otolith weight against the NIRS predicted age-class for 2009 ($\chi^2_{(2\ \mathrm{d.f.})} = 6.08$; $p = 0.048$); and (iii) the age-class distribution predicted from otolith weight against the NIRS predicted age-class distribution for 2006 ($\chi^2_{(2\ \mathrm{d.f.})} = 9.09$; $p = 0.011$).

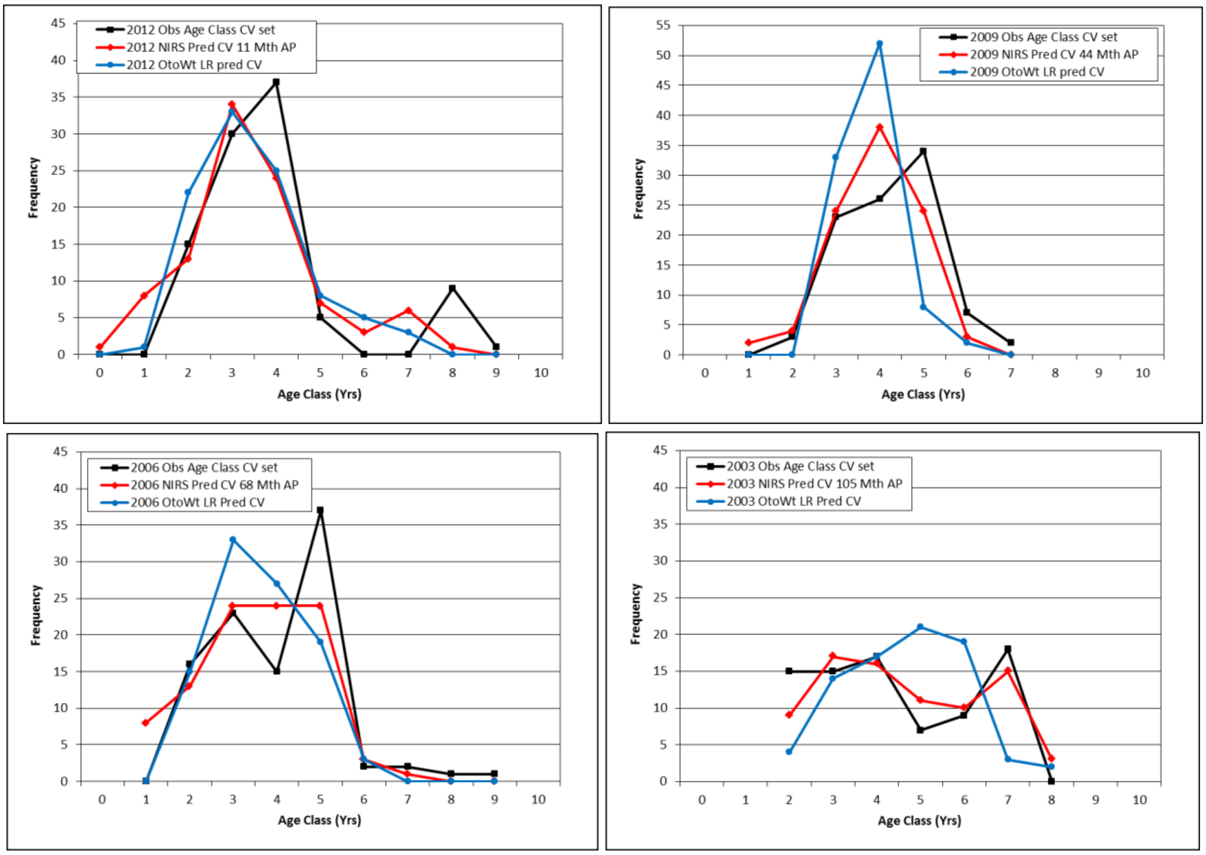


Figure 34. Observed and predicted age-class distributions for historic Fitzroy River Barramundi cross-validation sets

Results of fisheries performance measures from the analysis of ‘historic’ Fitzroy River Barramundi otoliths ≤120 months suggested reasonable efficacy of NIRS to estimate fish age, but substantial age-bias and poor between-method precision (i.e., IAPE) would benefit from further research.

Daly River Barramundi

Daly River Barramundi collected in July and September 2011 had variable length-at-age (Figure 35).

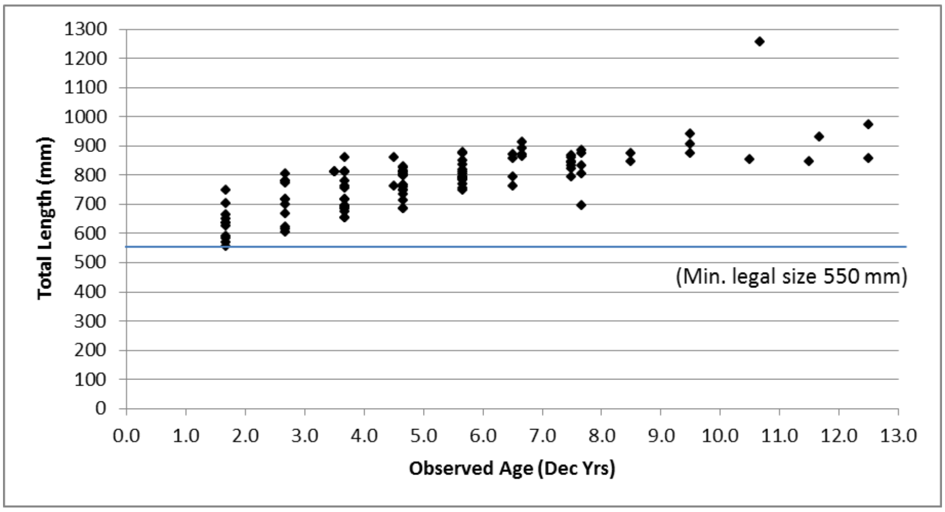


Figure 35. Length-at-age plot of the 2011 Daly River Barramundi otolith samples

NIRS predicted ages had better model performance than age predicted from otolith weight and total length (Table 14). The correlation coefficient (r) for the NIRS model was 0.90 and the slope of the NIRS models was 0.82. Age-class derived from NIRS predicted age was correct for 56% of the 2011 Daly River CV set, with 51% and 82% of individuals predicted to within 6 months and 12 months of their observed age respectively. The IAPE for NIRS age estimates was 7.8%.

Table 14. Performance of models predicting fish age using NIRS and regression analysis for 2011 Daly River Barramundi

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	% Correct			IAPE (%)
					Age-Class	+/- 6 Months	+/-12 Months	
2011 Cross-Validation	NIRS, LV=2	0.82	10.5	-0.10	56	51	82	7.8
N=95	LR OtoWt	0.69	13.7		43	55	66	9.8
Ages = 20 - 114 months	LR OtoWt+TL	0.69	13.7		43	46	63	9.6
SD = 24.8 months								

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; LV = Latent Variables; LR = Linear Regression

For any observed age, predicted age was variable (Figure 36). Age-bias plots indicated a slight pattern of bias in NIRS predicted ages for Daly River Barramundi (Figure 37), but this bias was less than the bias for age predicted from a linear regression of otolith weight (Figure 37).

Predicted age-class distributions were similar to the observed age-class distribution for Daly River Barramundi (Figure 38), with no significant differences between the observed and NIRS ($\chi^2_{(2\text{ d.f.})}$ = 0.76; p=0.685); observed and linear regression of otolith weight ($\chi^2_{(2\text{ d.f.})}$ = 1.71; p=0.426); or NIRS and linear regression of otolith weight ($\chi^2_{(2\text{ d.f.})}$ = 0.76; p=0.685).

The tendency for age-bias and issues with between method precision (i.e., IAPE = 7.8) suggests that further research is required.

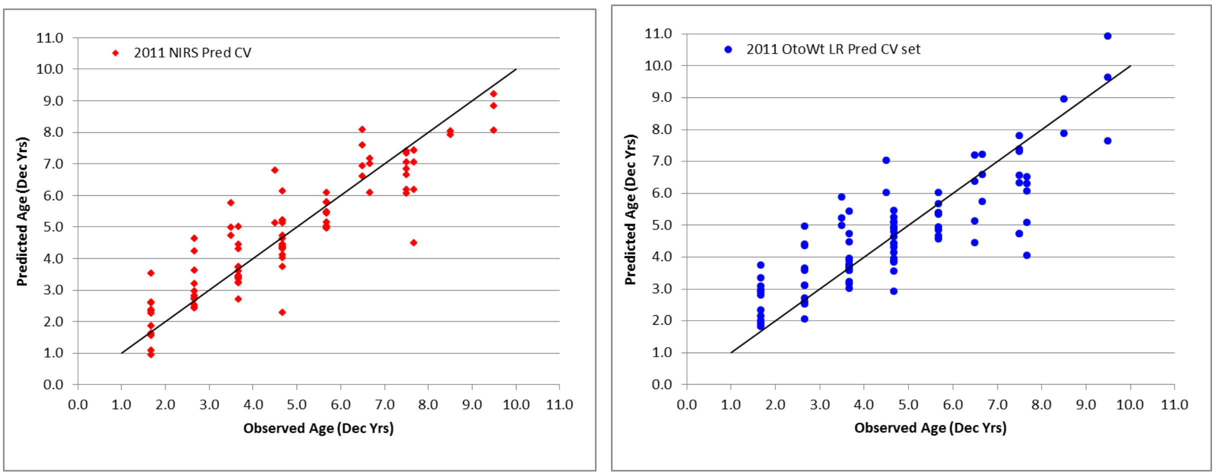


Figure 36. Observed age versus predicted age for 2011 Daly River Barramundi cross-validation set; solid black line indicates the 1:1 equivalence

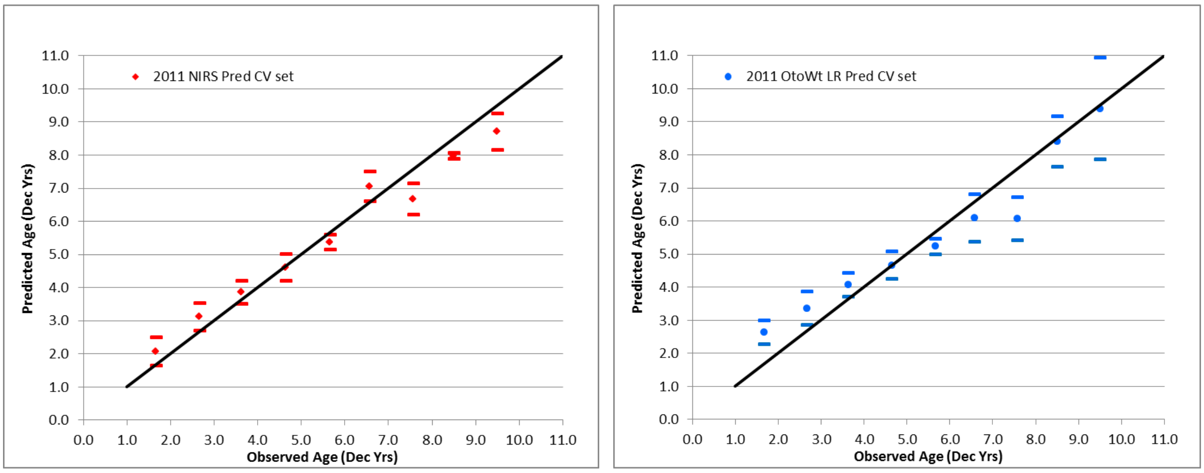


Figure 37. Age-bias plots for 2011 Daly River Barramundi cross-validation set; error markers are the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence

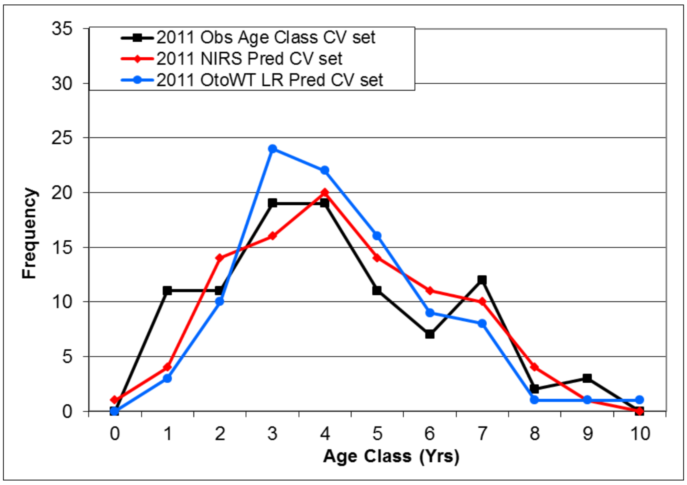


Figure 38. Observed and predicted age-class distributions for 2011 Daly River Barramundi

Snapper

Sunshine Coast Offshore Waters, Queensland

The 2012 samples of Snapper collected from the Queensland Sunshine Coast offshore waters had variability in length-at-age (Figure 39) typical for Snapper.

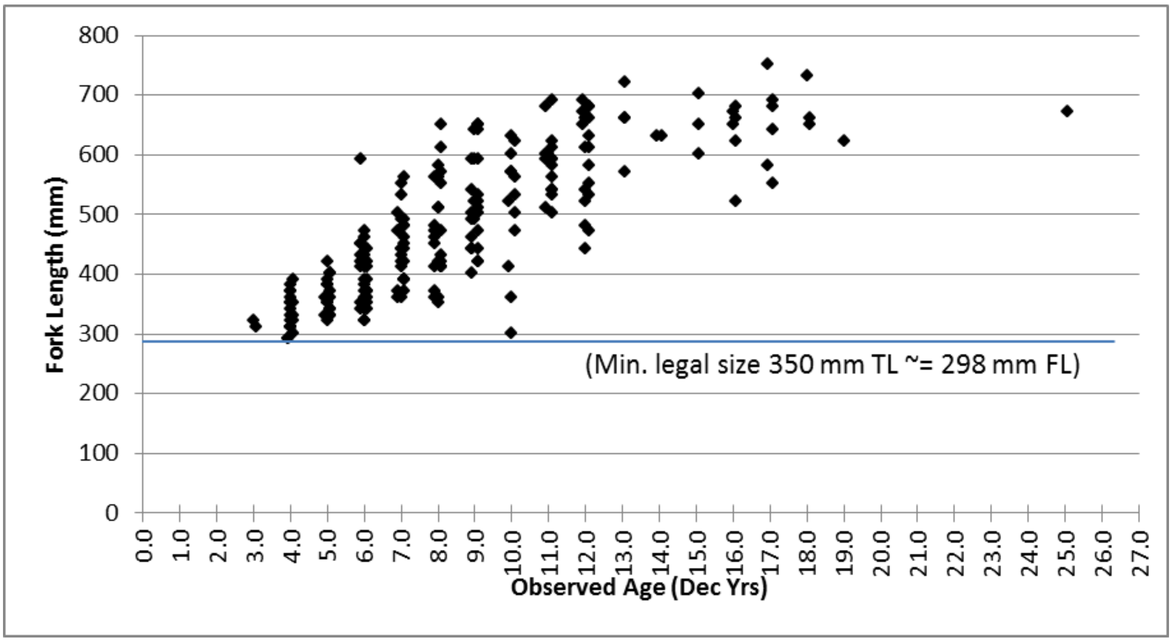


Figure 39. Length-at-age plot of 2012 Sunshine Coast Snapper otolith samples

Predicting Snapper age based on NIRS performed similarly to age predicted from otolith weight (or otolith weight plus fork length), with R^2's for validation sets between 0.69 and 0.79 and RMSE's ~15 months for NIRS predictions and ~17 months for otolith weight based predictions (Table 15). NIRS predicted age was converted to age-class by rounding down to the nearest integer and had slightly better performance than age-class based on otolith weight. The percentage of individuals predicted to within 6 or 12 months of their observed age was much lower for Sunshine Coast Snapper than for any of the Barramundi analyses. IAPE was 6.3% for the validation set based on a 3-month acquisition point (Table 15). Note that about 15% of the 2012 Sunshine Coast Snapper otoliths were broken or had major chips and were removed from the data sets used in the linear regression of otolith weight to predict age, shown by the number of outliers removed (Table 15).These samples were not removed from the NIRS calibration or validation sets. Therefore, the results for NIRS age prediction are not directly comparable to results for linear regression of otolith weight. However, each set is directly comparable to the observed age values. The inability to use broken or chipped otoliths in a regression of otolith weight in predicting fish age is another reason why otolith weight is not widely used to predict fish age.

Table 15. Performance of models predicting fish age using NIRS and regression analysis for 2012 Sunshine Coast Snapper

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Calibration N=100 Ages = 36 - 145 months SD = 32.3 months	NIRS 3 month AP, (1 OR) LV=3	0.79	14.8	-0.14	0.89	0.80	39	31	59	7.0
	LR OtoWt (21 OR)	0.77	15.2				34	32	62	7.0
	MLR OtoWt +FL (21 OR)	0.77	15.2				32	32	62	6.9
	Power Reg. OtoWt (21 OR)	0.78	14.9				38	33	53	6.7
2012 Validation N=91 Ages = 48 – 145 months SD = 26.9 months	NIRS 3 month AP, LV=3	0.69	14.9	2.28	0.83	0.71	35	32	54	6.3
	LR OtoWt (10 OR)	0.71	17.7				31	33	57	7.3
	MLR OtoWt +FL (10 OR)	0.71	17.7				31	33	57	7.2
	Power Reg. OtoWt (10 OR)	0.72	17.7				31	27	60	7.3
2012 Cross Validation N=94 Ages = 36 – 144 months SD = 27.9 months	NIRS 6 month AP, LV=2	0.70	15.4	-0.12	0.83	0.71	34	27	51	7.9
	NIRS 12 month AP, LV=2	0.71	15.2	-0.16	0.84	0.72	34	29	51	7.6
	NIRS 15 month AP, LV=2	0.70	15.4	-0.18	0.83	0.72	35	29	48	7.9

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables; OR = Outliers Removed; LR = Linear Regression; MLR = Multiple Linear Regression; Power Reg.: age = $a+b*OtoWt^c$

For any observed age, NIRS and power regression predicted age was variable (Figure 40). Age-bias plots indicated some bias was present in the NIRS predictions as well as the otolith weight predictions (Figure 41). Observed and predicted age-class distributions were similar (Figure 42), with no significant differences between the observed and NIRS predicted ($\chi^2_{(2\text{ d.f.})} = 1.08$; p=0.584); or between the observed and otolith weight predicted ($\chi^2_{(2\text{ d.f.})} = 3.56$; p=0.169).

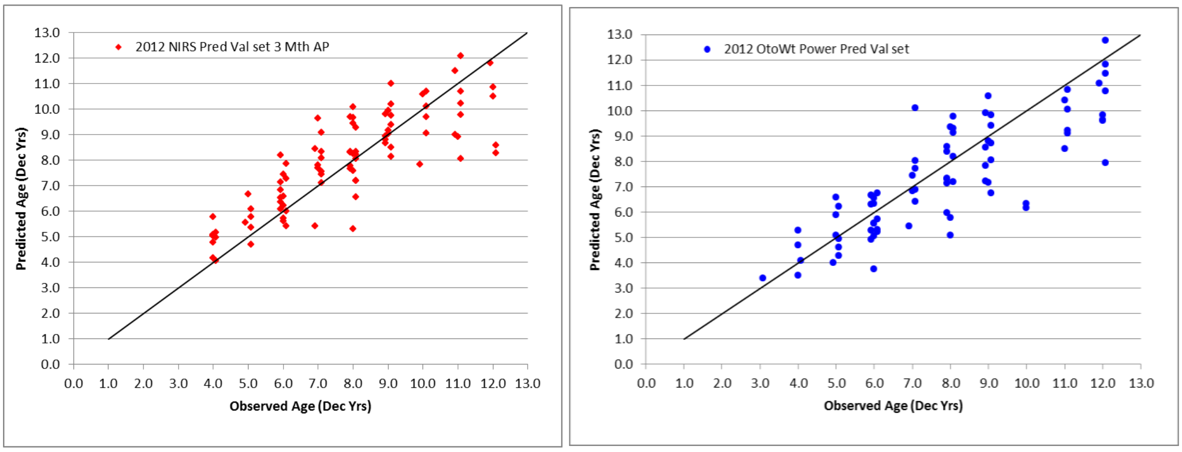


Figure 40. **Observed age versus predicted age for 2012 Sunshine Coast Snapper validation set; solid black line indicates the 1:1 equivalence**

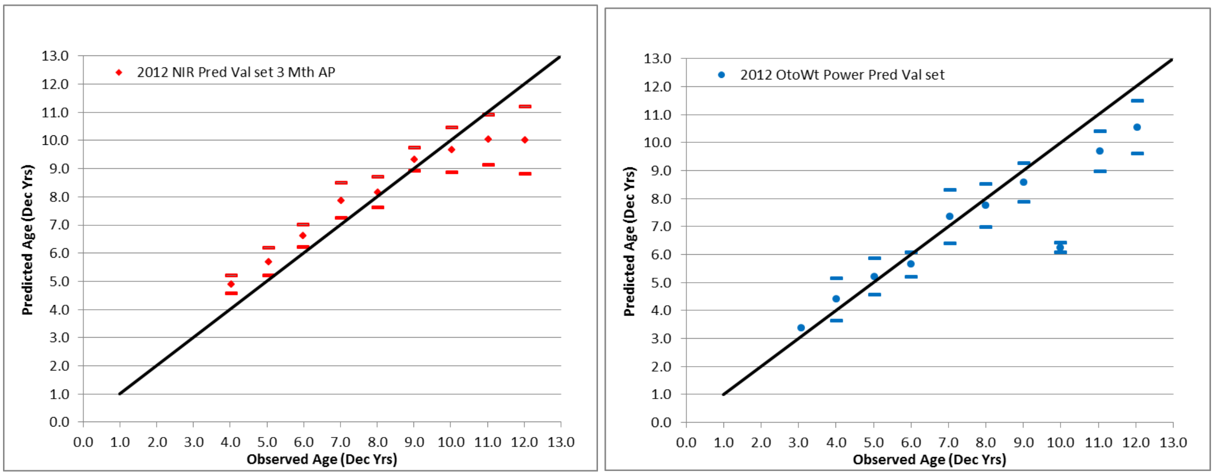


Figure 41. **Age-bias plots for 2012 Sunshine Coast Snapper validation set; error markers represent the 95% confidence interval about the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence**

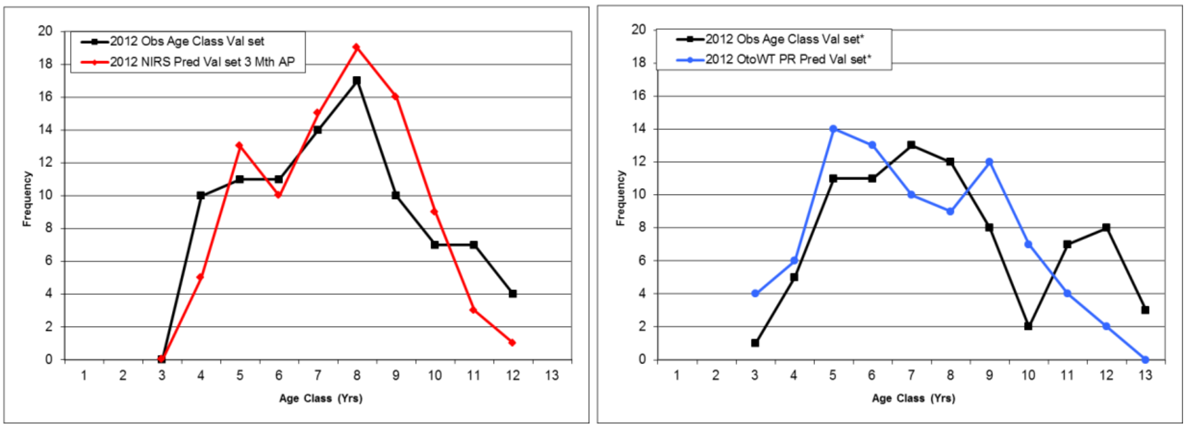


Figure 42. Observed and predicted age-class distributions for 2012 Sunshine Coast Snapper validation set

Despite having good NIRS model performance, the fisheries performance measures indicated poor efficacy for NIRS age estimates of Snapper <156 months. Of concern amongst fisheries end-users consulted by the project were the large RMSE's, the presence of age-bias and poor between-method precision (i.e., IAPE's ≥6.3%). Further research is required to understand the causes of these results and whether data transformations prior to analysis could ameliorate these issues.

Two additional historic sets of Snapper otoliths from the Sunshine Coast were also assessed. Otoliths were collected in June to August 2009 (n=97) and July to November 2006 (n=70) and showed typical variability in length-at-age (Figure 43).

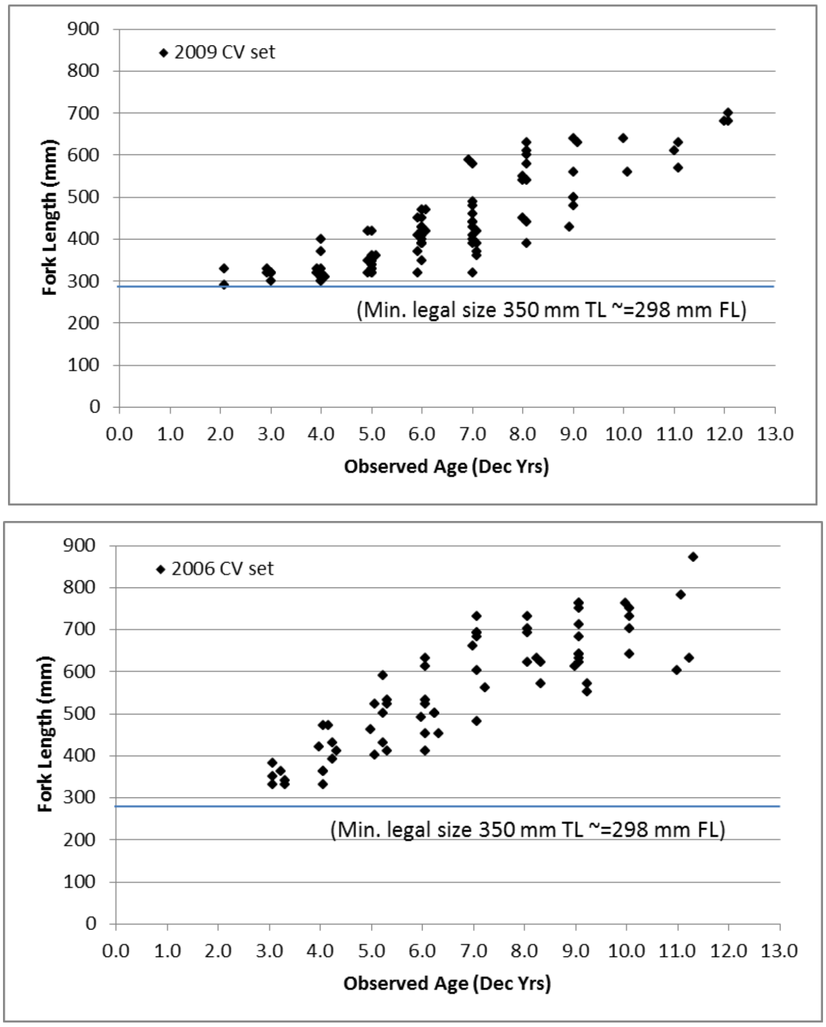


Figure 43. Length-at-age plots of historic Sunshine Coast Snapper otolith samples collected in: (a) 2009, n=97; and (b) 2006, n=70

These historic sets of Sunshine Coast Snapper otoliths were assessed for performance using cross-validation (CV) for individual collection years and calibration/validation sets for a combined collection year model. Model performance varied between collection years (Table 16). Age was predicted using a power regression of otolith weight (i.e., age = a+b*OtoWt^{c}) because this form of regression gave substantially higher R^{2}'s and lower RMSE's than linear regressions of otolith weight or otolith weight plus fork length. Across CV sets for historic sample years, age predicted from otolith weight had higher RMSE's, and lower percentages of individuals correctly assigned to age-class, or within 6 or 12 months of observed age than models based on NIRS (Table 16), except in 2006.

Table 16. Performance of models predicting fish age using NIRS and regression analysis for historic Sunshine Coast Snapper

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Cross-Validation N=94	NIRS 12 month AP, LV=2	0.71	15.2	-0.16	0.84	0.72	34	29	51	7.6
Ages = 64 - 144 months SD = 27.9 months	Power Reg. OtoWt (3 OR)	0.55	17.3				27	21	37	8.9
2009 Cross-Validation N=97	NIRS 51 month AP, LV=3	0.82	11.3	-0.16	0.91	0.83	35	51	72	5.9
Ages = 25 – 145 months SD = 26.9 months	Power Reg. OtoWt	0.79	12.				31	38	70	6.7
2006 Cross-Validation N=70 (1 OR)	NIRS 85 month AP, LV=3	0.80	12.6	0.10	0.89	0.81	19	32	62	6.7
Ages = 37 – 136 months SD = 27.8 months	Power Reg. OtoWt	0.75	14.0				22	36	68	7.0
All Yrs Calibration N=140	NIRS, LV=2	0.74	14.3	-0.02	0.86	0.74	33	31	60	7.0
Ages = 24 - 145 months SD = 27.8 months	Power Reg. OtoWt	0.70	18.2				25	31	57	9.9
All Yrs Validation N=123 (1 OR)	NIRS, LV=2	0.80	12.5	0.44	0.90	0.87	33	38	67	6.9
Ages = 24 - 144 months SD = 27.9 months	Power Reg. OtoWt	0.76	16.9				32	22	48	9.5

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; AP = Acquisition Point; LV = Latent Variables. OR = outliers removed; Power Reg.: age = $a+b*OtoWt^c$

For any observed age, NIRS predicted age and predicted ages from otolith weight were variable (Figure 44).

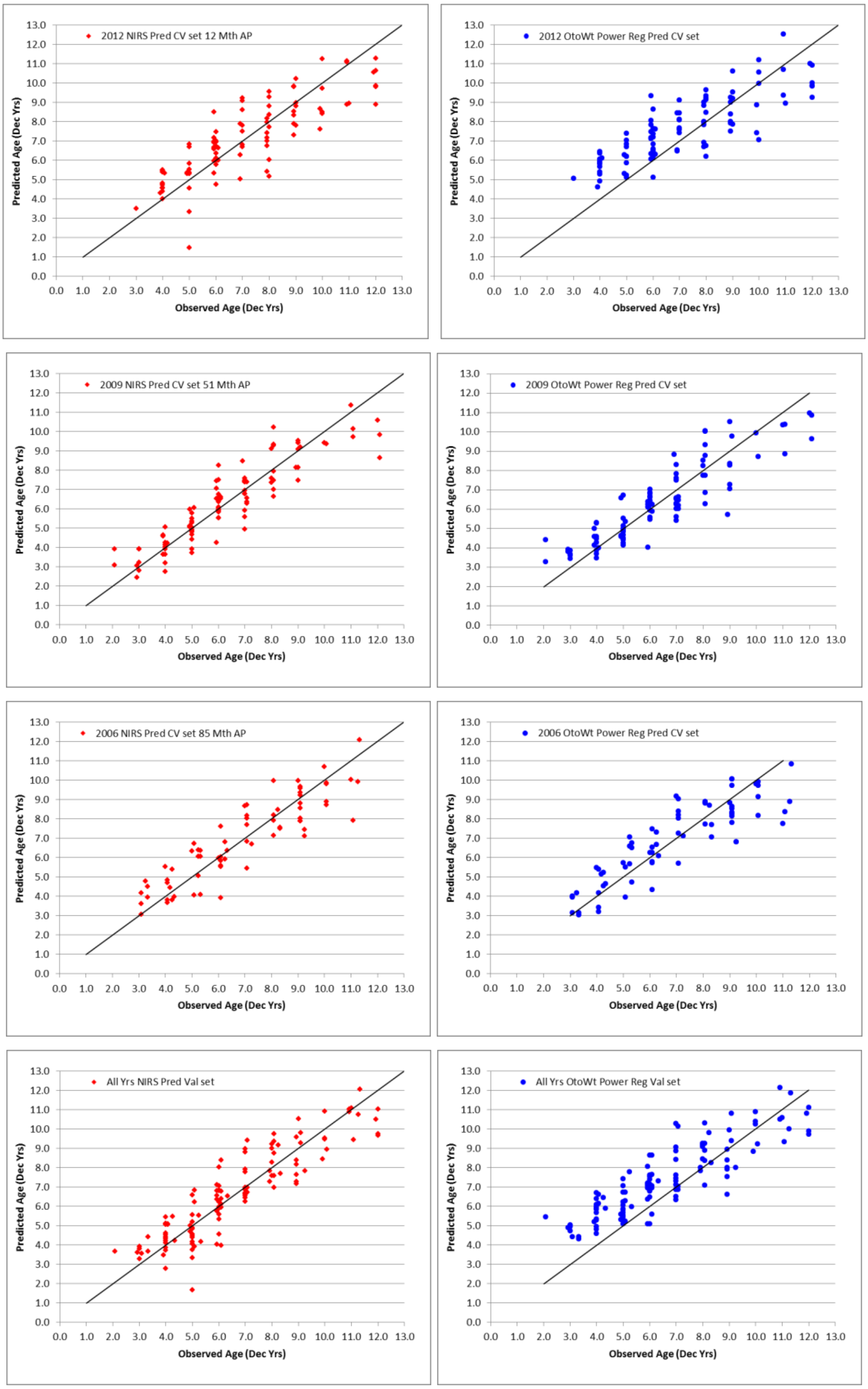


Figure 44. Observed age versus predicted age for historic Sunshine Coast Snapper; solid black line indicates the 1:1 equivalence

A small amount of age-bias was present in the NIRS predictions for Sunshine Coast Snapper (Figure 45), but was less than that observed for age predictions based on otolith weight.

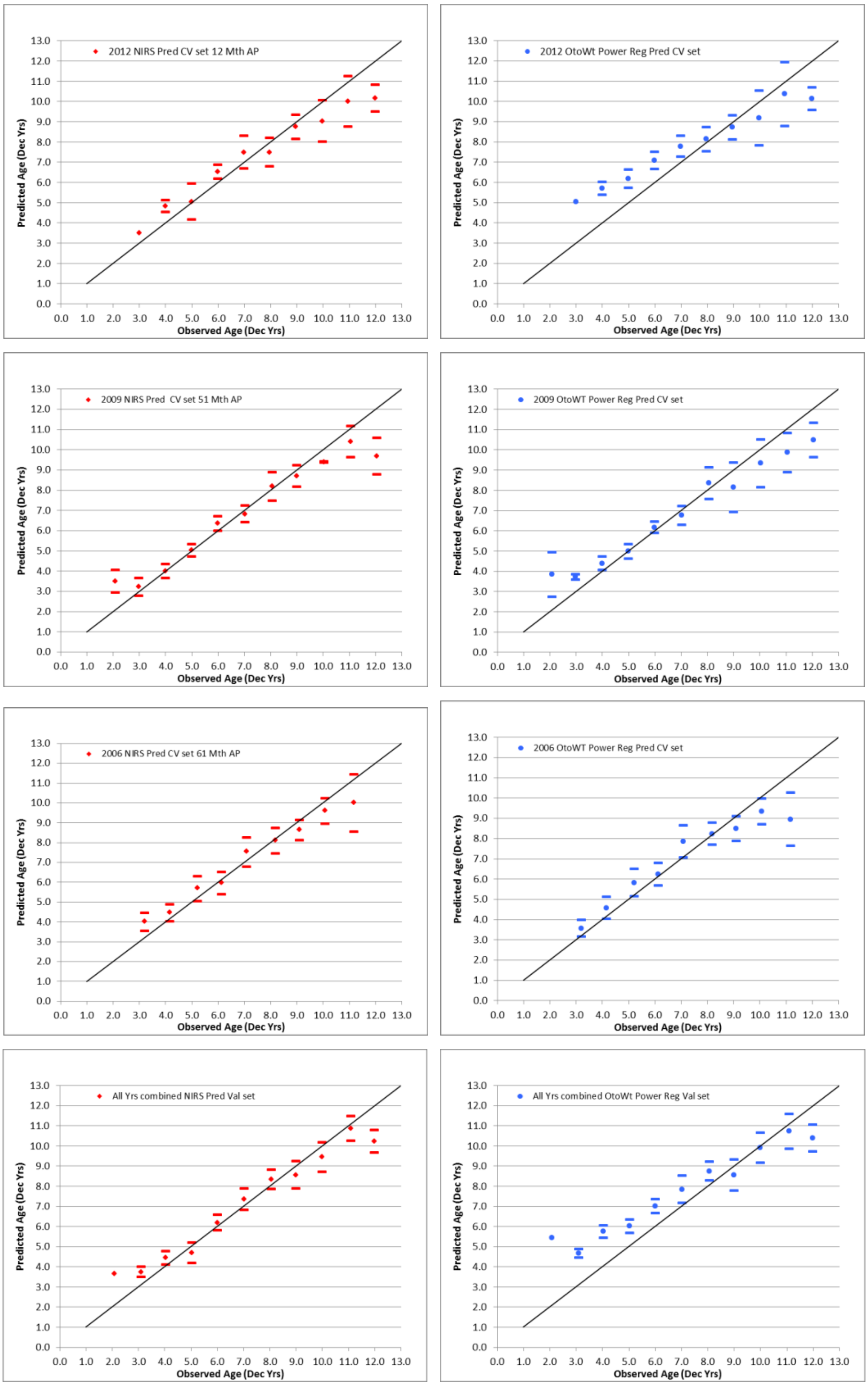


Figure 45. **Age-bias plots for historic Sunshine Coast Snapper. Error markers represent the 95% confidence interval around the mean predicted age per observed age group. Solid black line indicates the 1:1 equivalence.**

For the historic Sunshine Coast Snapper CV sets, observed and predicted age-class distributions were similar (Figure 46). In only three comparisons, were the distributions significantly different; all of which were the age-class distribution predicted from otolith weight against: (i) 2012 observed age class CV set ($\chi^2_{(2\text{ d.f.})}$ = 9.12; *p*=0.010); (ii) 2012 NIRS predicted age-class CV set ($\chi^2_{(2\text{ d.f.})}$ = 6.89; *p*=0.227); and (iii) all years combined NIRS predicted age-class validation set ($\chi^2_{(2\text{ d.f.})}$ = 9.68; *p*=0.008).

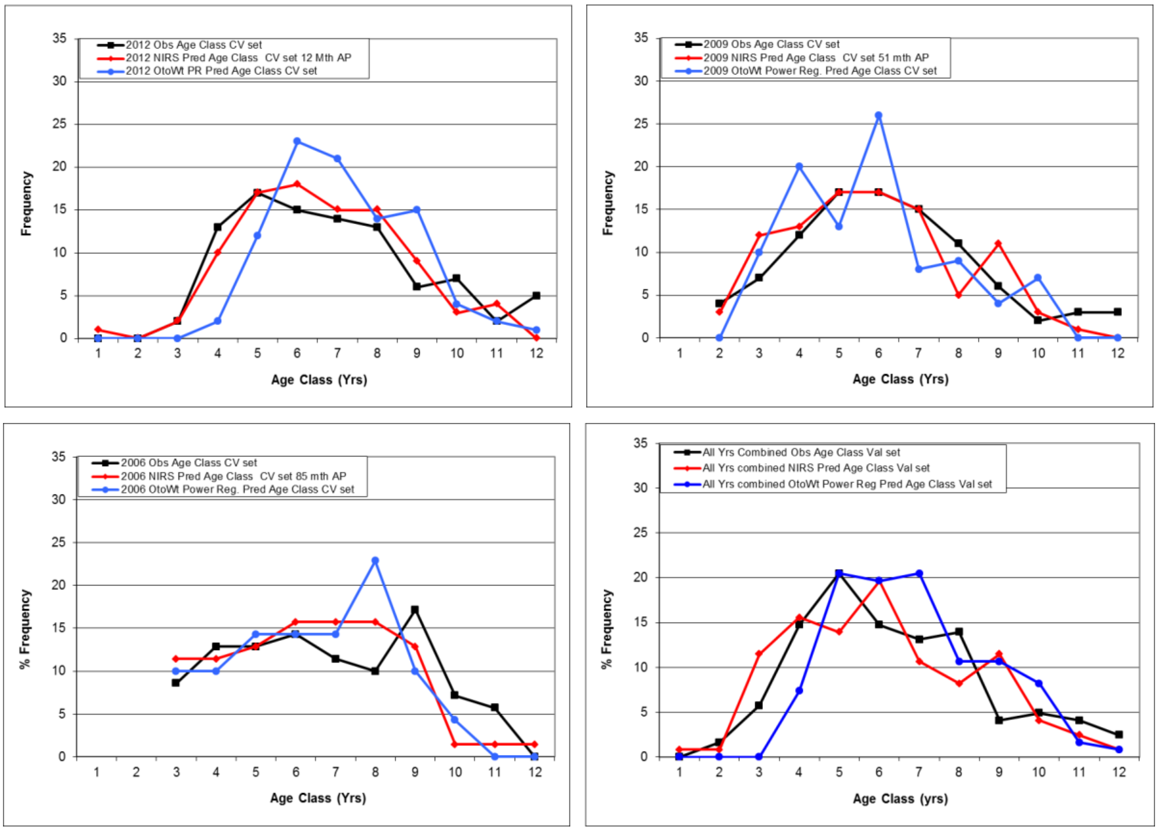


Figure 46. Observed and predicted age-class distributions for historic Sunshine Coast Snapper

Gulf St Vincent, South Australia

Samples of Snapper otoliths were collected from the Gulf St Vincent (GSV), South Australia, between January to May (Summer & Autumn) and June to September (Winter & Spring) 2012. Samples were pooled from the Northern and Southern GSV, although these areas are usually considered as separate stocks. Snapper in the GSV has variable length-at-age (Figure 47). Otoliths that were broken or with major chips were removed from the set using otolith weight to predict age but were not excluded from the NIRS data set unless the spectra obtained appeared atypical.

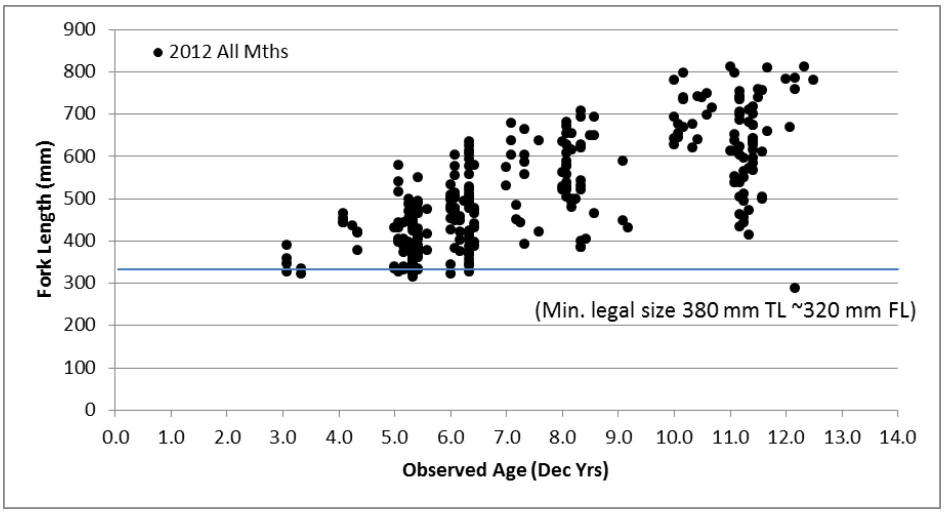


Figure 47. Length-at-age plot of 2012 Gulf St Vincent Snapper otolith samples, all months combined

For GSV Snapper, NIRS provided better estimates of fish age than regressions using otolith weight, with R^2's ≥0.85 and RMSE's ~11 months compared to R^2's of ~0.60 and RMSE's of >17 months respectively (Table 17). NIRS predicted age had better performance than age predicted from otoliths weight in terms of percentage age-class correct; percentage of individuals predicted to within 6 or 12 months of their observed age and between method precision (i.e., IAPE).

Table 17. Performance of models predicting fish age using NIRS and regression analysis for Gulf St Vincent Snapper

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	r	Slope	% Correct			IAPE (%)
							Age-Class	+/- 6 Months	+/- 12 Months	
2012 Summer &Autumn Calibration	NIRS, LV=2	0.85	10.4	-0.02	0.92	0.85	48	36	72	5.1
N=130 Ages = 37-148 months SD = 27.0 months	LR OtoWt, (3 OR)	0.68	14.8				33	40	63	6.7
2012 Summer & Autumn Validation	NIRS, LV=2	0.85	11.2	-1.18	0.92	0.84	51	52	77	4.6
N=115 Ages = 37–145 months SD = 29.2 months	LR OtoWt, (1 OR)	0.57	19.7				31	37	56	8.0
2012 Winter & Spring Cross-Validation	NIRS, LV=2	0.90	10.6	0.01	0.95	0.91	46	42	77	4.1
N=69 Ages = 65–150 months SD = 31.2 months	LR OtoWt	0.68	17.92				32	30	58	6.8
2012 All months Calibration	NIRS, LV=2	0.89	10.2	-0.02	0.95	0.90	50	43	74	4.8
N=148 Ages = 37-150 months SD = 32.0 months	LR OtoWt	0.64	18.6				32	34	51	8.0
2012 All months Validation	NIRS, LV=2	0.84	11.0	-1.40	0.92	0.83	52	46	77	4.7
N=164 Ages = 40-146 months SD = 28.0 months	LR OtoWt	0.63	16.9				41	37	62	6.8
2009 All months Cross-Validation	NIRS, LV=2, (2 OR)	0.92	11.40	-0.15	0.96	0.92	36	43	73	6.7
N=85 Ages = 25-153 months SD = 39.8 months	LR OtoWt, (21 OR)	0.93	10.32				44	52	78	5.2

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; LV = Latent Variables; LR = Linear Regression; OR = Outliers Removed

For any observed age, predicted age was variable (Figure 48). There was no consistent pattern of bias in the NIRS age predictions (Figure 49), but there was a slight pattern of bias in otolith weight based age predictions.

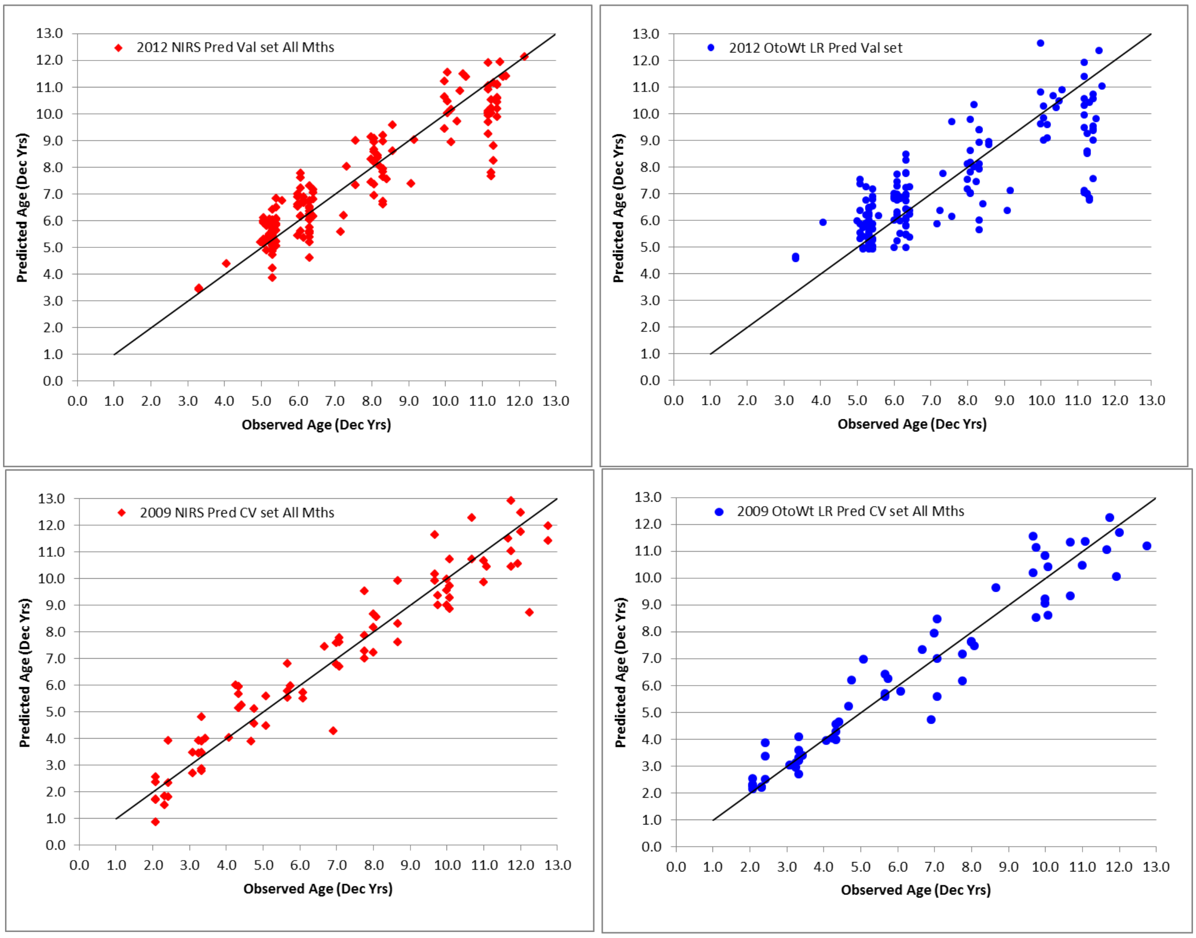


Figure 48. Observed age versus predicted age for Gulf St Vincent Snapper; solid black line indicates the 1:1 equivalence

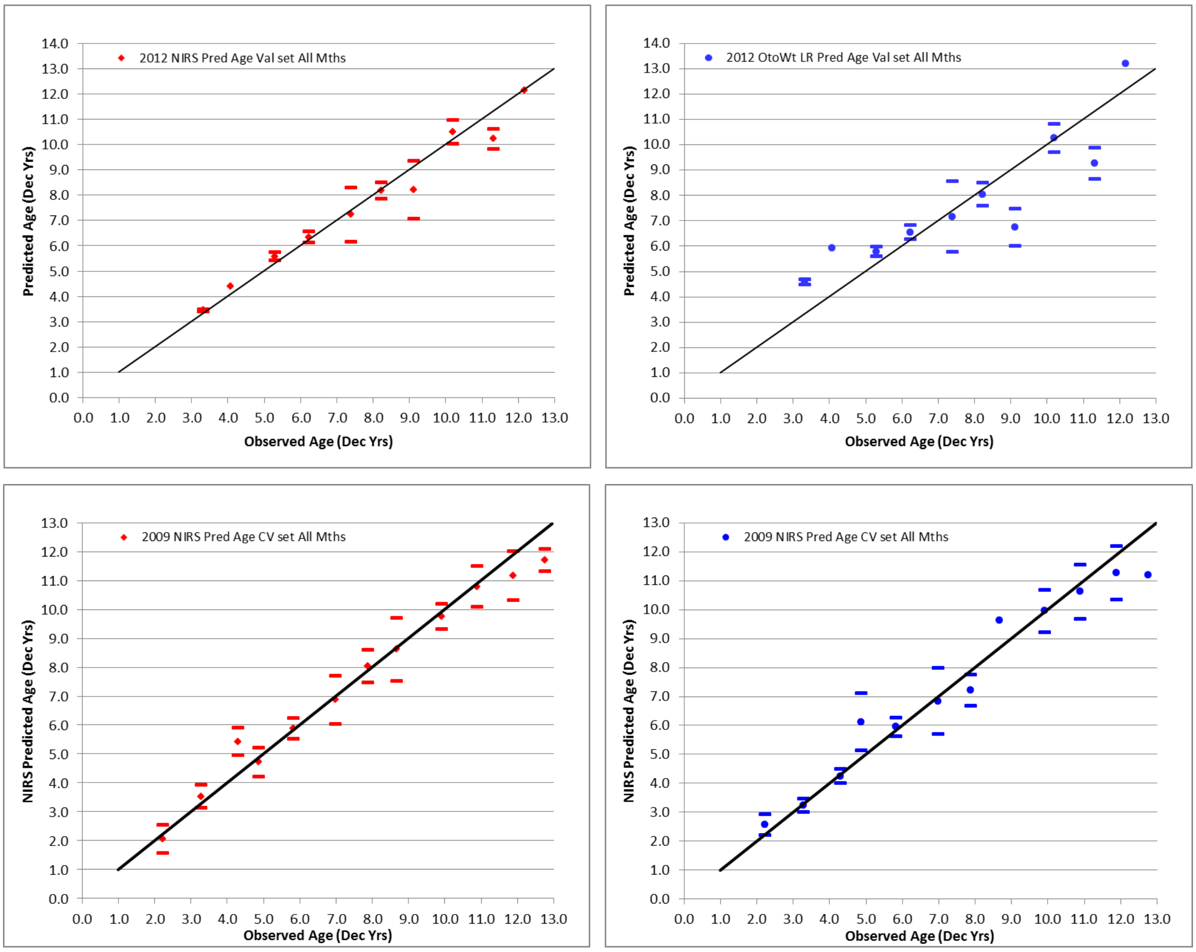


Figure 49. Age-bias plots Gulf St Vincent Snapper; error markers represent the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence

None of the predicted age-class distributions were significantly different from the observed age-class distributions (Figure 50), i.e., recorded *p* values were all >0.05.

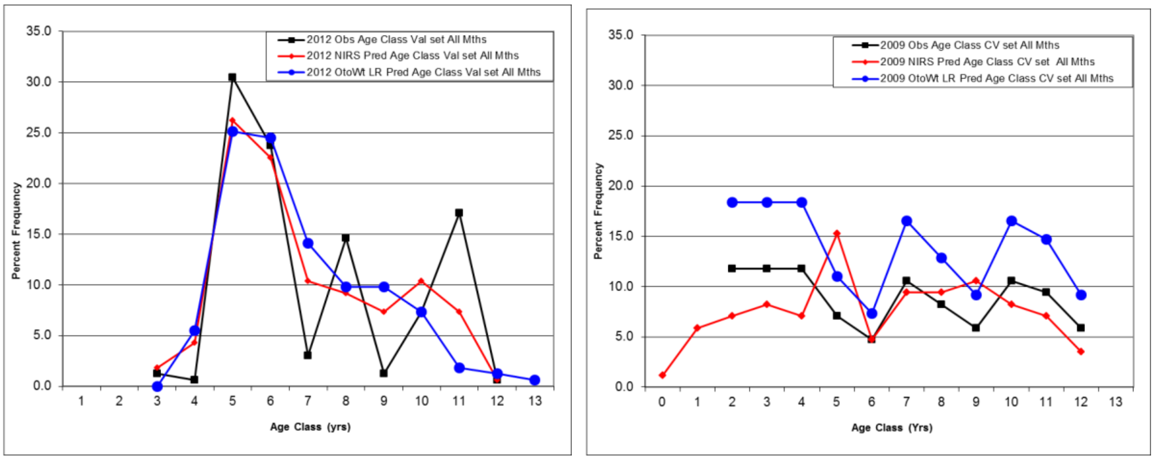


Figure 50. Observed and predicted age-class distributions for Gulf St Vincent Snapper

Mid-West Coast, Western Australia

Snapper otoliths were collected from the Mid-West Coast of Western Australia in 2010 and showed variable length-at-age (Figure 51).

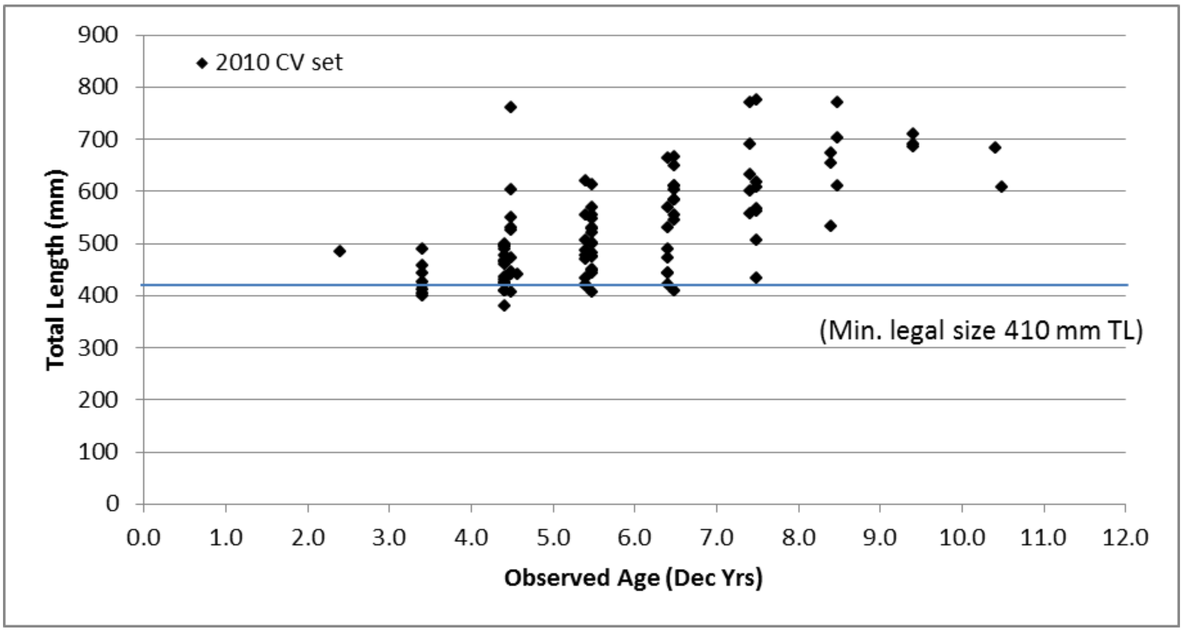


Figure 51. Length-at-age plot of 2010 Mid-West Coast Snapper otolith samples

For WA Snapper, estimates of fish age based on NIRS were similar in performance to estimates of fish age based on regressions using otoliths weight (plus fork length). For the NIRS model, $R^2 = 0.58$, RMSE = 13.0 months, the correlation coefficient (r) = 0.76 and the slope = 0.59. For the otolith weight model, $R^2 = 0.48$ and RMSE = 14.4 months (Table 18). Models based on NIRS and otolith weight had similar performance when assessed for their ability to provide a basis for correctly allocated age-class, or correctly predicted to within 6 or 12 months of their observed age (Table 18).

Table 18. Performance of models predicting fish age using NIRS and regression analysis for 2010 Mid-West Coast Snapper

Data Set	Model	R^2	RMSE (Months)	Bias (Months)	% Correct			IAPE (%)
					Age-Class	+/- 6 Months	+/- 12 Months	
2010 Cross-Validation N=100	NIRS, LV=3	0.58	13.0	0.03	42	43	69	7.2
Age = 29-126 months SD = 18.9 months	LR OtoWt	0.48	14.4		40	41	66	8.3

RMSE = Root Mean Square Error; IAPE = Index of Average Percent Error; LV = Latent Variables; LR = Linear Regression

For any observed age, predicted age was variable (Figure 52). Age-bias plots for Mid-West Coast WA Snapper indicated consistent patterns of bias in predicted age (Figure 53). Despite the patterns of bias, there were no significant differences between the observed age-class distribution and either of the predicted age-class distributions (Figure 54), i.e., recorded *p*-values were >0.05.

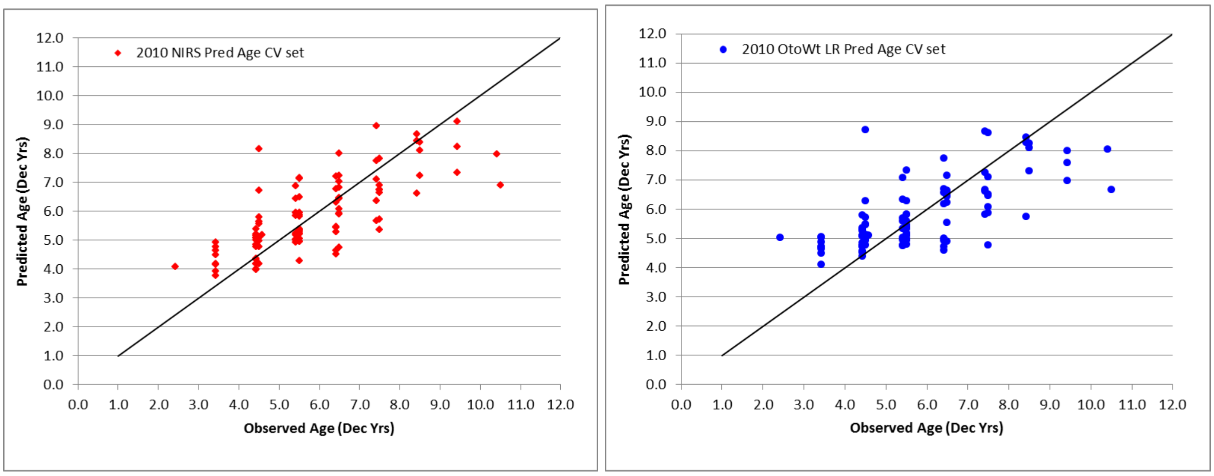


Figure 52. **Observed age versus predicted age for 2010 Mid-West Coast Snapper; solid black line indicates the 1:1 equivalence**

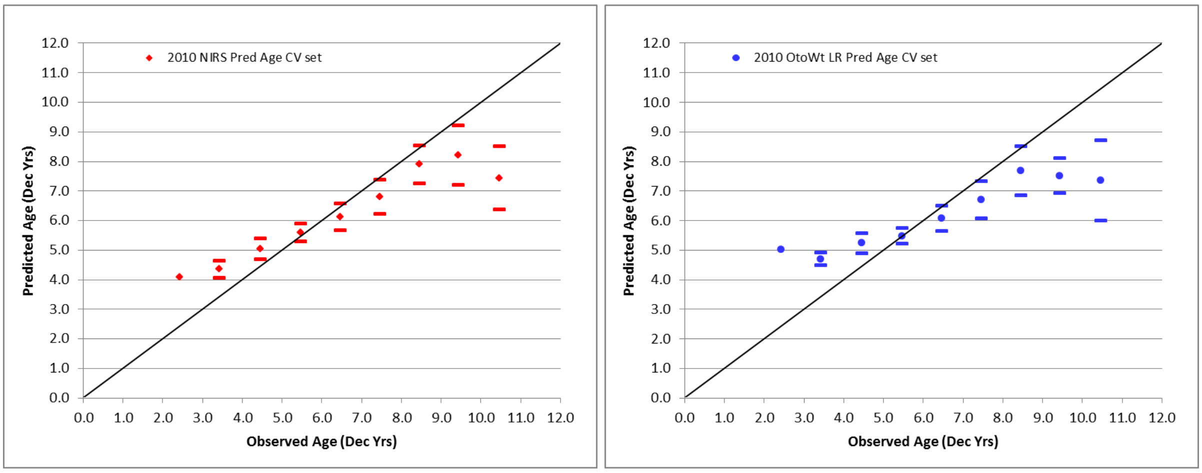


Figure 53. **Age-bias plots of 2010 Mid-West Coast Snapper; error markers represent the 95% confidence interval around the mean predicted age per observed age group; solid black line indicates the 1:1 equivalence**

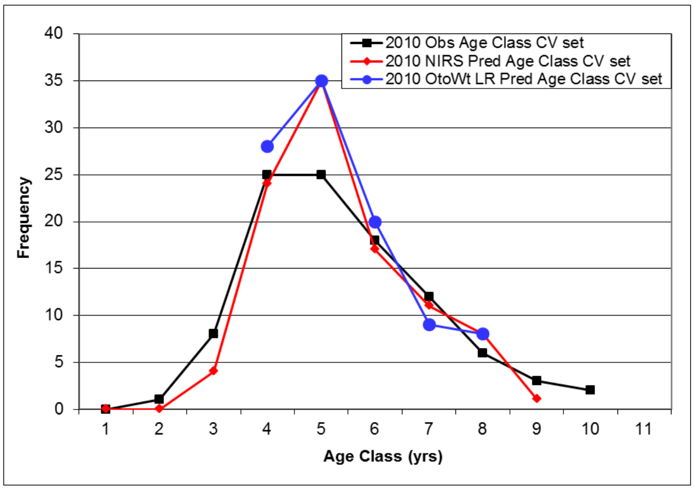


Figure 54. **Observed and predicted age-class distributions for 2010 Mid-West Coast Snapper cross-validation sets**

Objective 4. Cost-effectiveness of ageing fish using NIRS

Otolith collection and processing costs

Each state agency supplied (where possible) estimated costs per otolith for standard ageing (Table 19). Not surprisingly, otoliths were more expensive to collect than process (Table 19), being most expensive to collect from remote areas; a general finding already identified by Craine *et al*. (2009). Although Craine *et al*. (2009) defined the concept of Kolmogorov-Smirnov precision and used this to determine minimum sample sizes required for proxy ageing methods to achieve set precision values, the otolith samples used in the current project were insufficient to employ this method. Instead, we use a 2012 collection from the one of the Barramundi stocks in Queensland as a case study of the potential cost involved in ageing fish using standard methods compared to the hypothetical application of NIRS technology.

Table 19. Costs of otolith collection and processing

Species	Location	Cost Per Otolith*	
		Collection	Processing
Barramundi	Gulf of Carpentaria, Queensland	$85 ($70 - $107)	$14.6
	Fitzroy River, Queensland	$47	$22
	NT Daly River, recreational fishers	$18	$32
	NT Daly River, commercial fishers	$30	$25
Snapper	Queensland	Not estimable	$21
	SA Gulf St Vincent	$36	$32
	WA Ocean Gascoyne	$130	
	WA Perth Metro	~$30	

* Costing's supplied by project collaborators and project workshop participants, see Appendix 1

In addition to supplying the above costs of otolith collection and processing, each collaborating state agency also commented on the purpose of ageing fish. For Queensland, South Australia and Western Australia, the primary purpose of ageing fish was to generate annual age-length keys for each regional stock of a species. These age-length keys are used to generate annual age frequencies/proportions at age/population age structure for each regional stock. Further analyses included annual estimates of mortality (Z), based on cross-sectional or longitudinal analysis of catch curves; and calculation of recruitment histories based on residuals from catch curve analysis; often to identify strong (or weak) year-classes. The level of assessment of Barramundi and Snapper stocks by collaborating agencies varied; ranging from semi-quantitative/qualitative stock status assessments (see www.fish.gov.au) conducted on an annual basis, to highly quantitative age-structured fishery assessment models constructed at three year intervals (Snapper South Australia) or more intermittently (Snapper Queensland).

Cost-effectiveness

It was difficult to evaluate the cost-effectiveness of ageing fish by NIRS, because the models developed in the current project were intended as 'proof of concept', were generally focused on generic waveband selection rather than optimising the model for regions or seasons, and it difficult to predict what level of NIRS calibration model maintenance might be applicable. Therefore, it was difficult to "develop optimised fish sampling regimes with respect to cost" as per objective 4 in the original project proposal. In addition, several of the fisheries end-users commented that they would not really consider cost effectiveness until further work (such as the pilot study of NIRS applied to the 2012 southern Gulf of Carpentaria Barramundi stock) resolved some of the issues identified, such as the benefits of split age models, transformation of age data prior to analysis to address issues of age-bias and how age-length variability could be incorporated into NIRS model development and subsequent NIRS model maintenance.

However, costs in applying NIRS to predict the age of fish otolith samples involves:
(i) otolith collection costs (same as traditional method);
(ii) processing costs of otoliths aged by standard visual assessment of otolith thin sections (same as traditional method) for use in NIRS calibration model development or maintenance;
(iii) NIR spectroscopy equipment costs (outright purchase, lease or outsource samples to existing laboratory with relevant NIRS equipment and expertise);
(iv) labour costs for NIR spectra collection from otolith samples;
(v) labour costs for development of the NIRS calibration model (i.e., chemometric analysis);
(vi) labour costs for NIR spectra collection from unknown otoliths and concurrent age prediction for these otoliths from applying the NIR calibration model;
(vii) labour costs for ongoing maintenance of NIR calibration model; and
(viii) ongoing costs of NIR spectroscopy equipment parts and service(s).

The costs associated with applying NIRS to age fish were separated into two general components: (i) NIRS model development (i.e., chemometric analysis); and (ii) routine application. The latter (i.e., routine application) is dealt with first, because it entailed fewer assumptions than the former (i.e., what would model development entail). We used a Queensland Barramundi sample as a case study to illustrate hypothetical cost-efficiencies.

Case study: routine application of NIRS to Barramundi in Queensland

Background information

Barramundi has seven regional stocks in Queensland, from which three are specifically sampled by Fisheries Queensland to "*collect representative length, age and sex data from the recreational and commercial catch annually*" (Fisheries Queensland 2010). Barramundi sampling is well-documented; with specific sampling design and targets (e.g., to annually sample 50 catches from the three regional stocks that are monitored). Estimated ages of Queensland Barramundi are used as part of the sub-sampling method of deriving annual age-length keys (ALK) for each regional stock that can be applied to fish measured within a stock and then upscaled (by the proportion of catch sampled) to estimate the annual age-structure of the Barramundi populations in three key areas of Queensland. These data are used as indicators of the stock and will potentially be incorporated in to highly quantitative stock assessment models.

Barramundi are sampled from a variety of sources (commercial fishers, recreational fishers, commercial seafood processors), with individual fish (and otoliths) having different post-mortem handling. Some otoliths are extracted from fresh fish; others are extracted from fish chilled for up to 7 days in ice brine, while others are extracted from frozen (and thawed) frames. The number of Barramundi sampled by Fisheries Queensland varies between years, but in 2012 across all regional stocks, 5686 Barramundi were measured for total length, while 1984 otoliths were aged by standard methods. Otoliths are selected for collection from length-measured Barramundi based on a target of collecting 20 otoliths per 10 mm (total) length class in each regional stock each year. There is a collection limit of five otoliths per length class from any individual catch to ensure collected otoliths are representative across catches.

Collection costs for Barramundi otoliths vary across the state, ranging from~ $50 to ~$110 per otolith (Table 19), reflecting the costs of accessing fish from areas close to major population centres compared to very remote areas, as well as the variable number of otoliths that are collected on any particular sampling trip. Processing costs for Barramundi were ~$15 per otolith, noting that

Queensland Barramundi are processed slightly differently to the standard ageing method, as five otoliths are blocked concurrently in a single block of resin and sectioned at the same time.

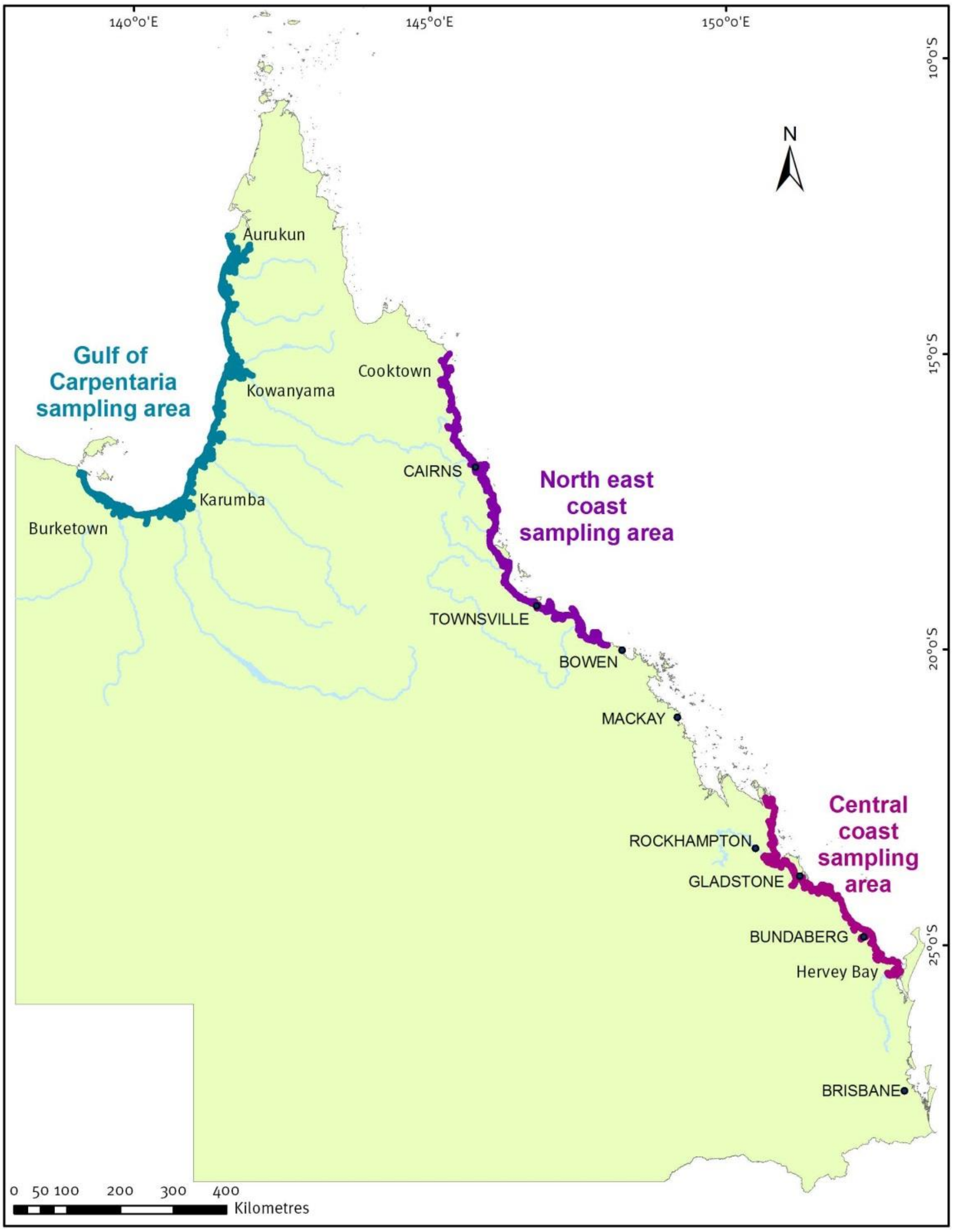


Figure 55. Barramundi stocks monitored by Fisheries Queensland as part of the Long Term Monitoring Program (from Fisheries Queensland 2010)

Routine application – 2012 southern Gulf of Carpentaria Barramundi

Hypothetical running costs were estimated for the routine application of NIRS age prediction models to the 2012 collection of Barramundi otoliths from the southern Gulf of Carpentaria stock (Table 20) and assumed acceptable levels of accuracy/efficacy of NIRS predictions to meet the data requirements of Fisheries Queensland. Costs (in labour hours and $) were estimated for two scenarios:

(i) a ‘best-case’, where the 2012 sample of otoliths was similar (in variability) to the that seen during NIRS calibration model development; therefore requiring only 10% of the annual otolith sample to be added to the NIRS calibration model for ‘model maintenance’; and
(ii) a ‘worse-case’, where the 2012 sample of otoliths was dissimilar (in variability) to that seen during NIRS calibration model development, therefore requiring 50% of the annual otolith sample to be added to the NIRS calibration model for ‘model maintenance’.

The number of otoliths that would need to be added to the calibration model (as part of NIRS maintenance) will vary between years, depending on the variability within the otoliths collected in any given year. How many otoliths need to be incorporated as part of NIRS model maintenance can be determined when the annual sample is scanned. Selecting samples for inclusion into a (global) calibration is often based on two statistical techniques, the ‘standard H’ statistic (Mahalanobis distance or Global H) and the ‘Neighbourhood H’ (NH) technique (see Shenk *et al.* (2001) for details of their application).

The hypothetical running costs assume that: it costs $14.6 per otolith to process for visual age assessment (Table 19), taking 0.232 hours of labour (in total) per otolith (split across several days); 100 otoliths can be scanned per hour for NIR spectra; and with labour costs (per hour including on-costs) of $125 for fisheries expertise and $150 for chemometric expertise.

Under the best-case scenario (i.e., 10% of annual otolith samples are used for NIR model maintenance and therefore require standard ageing), NIRS ageing of Barramundi was almost five times faster than the standard fish ageing technique. Under the ‘worst-case scenario’, where 50% of annual otolith samples are used for NIRS model maintenance, NIRS ageing was about 1.5 times faster than the standard fish ageing technique. Labour time efficiencies achieved by NIRS ageing of fish translate into significant cost savings (Table 20).

Table 20. Hypothetical running costs for NIRS age prediction applied to the 2012 collection of otoliths for the southern Gulf of Carpentaria Barramundi stock

Method	Otoliths Collected (n)	Otoliths Sectioned and Visually Assessed		NIR Scanning Labour (Hrs) @$125/hr	NIR model Maintenance Labour (Hrs) @$150/hr	QA* Visual Assessment (Hrs) @$125/hr	Total Hrs	Costs ($)
		(n)	(Hrs) @$125/hr					
Standard	605	605	140.4			20.3	160.7	$20,086
NIRS Add 10% (Best-Case)	605	61	14.0	6.05	2	13.3	35.4	$4,473
NIRS Add 50% (Worst-Case)	605	303	70.2	6.05	3	20.3	99.5	$12,516

* QA visual assessment: [(competency in fish ageing software program (FAS) = visual age estimate of 200 otoliths from reference collection) + (a re-read from the current years sample of ‘otoliths sectioned’, normally 150 to 200 otoliths)], where for Barramundi, it takes (on average) 3.05 minutes to visually age an otolith. The competency part of the QA of visual assessment happens regardless of the method or annual sample size.

NIRS calibration model development

It is difficult to forecast what is required (in terms of sample number, replicate years, labour costs and therefore dollar costs) to develop an NIRS calibration model to a level where ‘routine application’ is valid. In general, the development of the initial NIRS calibration model (including validation) requires substantial samples to incorporate the majority of the biological variability that occurs within the material of interest. As such, building a robust calibration model is a time-consuming and laborious procedure, and together with the complexity in the choice of data treatment are the main disadvantages of NIR spectroscopy.

The grain industry, which has been using NIRS technology for several decades, uses five to seven years of samples to build calibration models that incorporate sufficient between-year variability for models to be considered “robust” and thereafter just require model maintenance / ‘routine application’. NIRS calibration model development would need to include otolith samples that encompass all aspects of variability in NIR spectra with fish age over the following parameters: fish length; spatial location; age range; within-year and between-year variability; and differences post-mortem handling of otoliths (e.g., fresh fish, stored chilled fish, frozen and thawed fish).

The current project included 10 otoliths (where available) from each age-class for each species from each season or year from each geographic region to develop a calibration model. It is unknown whether this is sufficient to encompass all aspects of variability in NIR spectra with fish age. Application of the 2012 Archer River NIRS calibration model as part of a full pilot of the 2012 southern Gulf of Carpentaria Barramundi age and length data would assist in determining appropriate samples sizes for calibration models as applied to full scale fisheries data sets.

Procedures for calibration model development and maintenance are outlined in ‘Guidelines for the application of Near Infrared (NIR) Spectroscopy for predicting the age of whole dried otoliths’ (see Appendix 4).

NIR spectroscopy equipment costs

NIR spectrophotometer

Approximate costs of an ‘off the shelf’ NIR spectrophotometer range from $45,000 to $85,000 plus depending on wavelength range and purchased accessories (e.g., 30 position sample wheel). The Bruker Multi-Purpose Analyser (MPA - Figure 1) used in the current project costs about $80,000 plus and has the capability of transmission and reflectance modes for liquids and solids (see http://www.bruker.com/products/infrared-near-infrared-and-raman-spectroscopy/ft-nir/mpa/overview.html). A Bruker Tango-R costs $65,000+ but only has reflectance mode capabilities (see http://www.bruker.com/products/infrared-near-infrared-and-raman-spectroscopy/ft-nir/tango/overview.html). Cheaper bench top and hand held alternatives do exist and could potentially be used to age fish. Alternatively, a custom built NIR spectrophotometer could be used and cost from $15,000 onwards.

Spectrophotometers have a long service life (i.e., 20 plus years), but as with most scientific equipment, there are costs associated with yearly maintenance. Ongoing service costs vary depending on the make and model of the spectrophotometer and its location. Spectrophotometers require very few parts or consumables generally; except for an NIR light source and possibly a laser source. Both of these parts have thousands of hours in operating time before requiring replacement.

Chemometric software

Chemometric software is used to analyse the spectral data, develop the calibration model and predict future unknown samples. Many spetrophometers come bundled with chemometric software for example: OPUS comes with Bruker (http://www.bruker.com); WinISI with FOSS (http://www.foss.dk); VISION with Metrohm NIRSystems (http://metrohm-nirsystems.com); NIRCal with BUCHI (http://www.buchi.com), PICS with Perten (http://www.perten.com); and CalStar, UCal with Unity Scientific (http://www.unityscientific.com).

Alternatively, standalone multivariate chemometric software packages, such as ‘The Unscrambler’ used in the current project (http://www.camo.com), can be used for either quantitative or qualitative calibration model development. Standalone software packages include: GRAMS/AI + PLSPlus/IQ – (https://www.thermo.com); SIMCA-P+ (http://www.umetrics.com); PLS Toolbox, Solo (http://www.eigenvector.com); Pirouette (http://www.infometrix.com); SL Calibration Workshop (http://www.sensologic.com); AnalyzeIQ (http://www.AnalyzeIQ.com); and Matlab (http://www.mathworks.com.au/products/matlab/).

Costs vary greatly between packages, number of licences and uses (i.e., research, academic, commercial), with rough estimates of approximately $500 to $5000 plus. The choice of using bundled or standalone software often comes down to purpose, experience and personal/agency preference.

Overall, there are significant costs associated with an agency setting up a facility to employ NIRS to age fish (Table 21). These costs may be reasonable if applying NIRS to multiple species where large number of otoliths are collected and processed annually. Alternatively, it may be more feasible for a fisheries agency to out-source predicting fish age based on NIRS to groups like the Rapid Assessment Unit (collaboration between DAF and James Cook University) located in Cairns or a dedicated fish ageing consultancy like Fish Ageing Services.

Table 21. Summary of costs and variables associated with NIR spectroscopy to estimate fish age from otoliths

Description	**Summary:**
Spectrometer: • Purchase price range. • Depreciation costs 10 - 20 year life span. • Servicing - varies between instruments and locations for technician visit. • Parts/consumables - instrument specific.	• $15,000 to $85,000. • $1,500 - $750 to $8,500 - $4,250/year. • Off the shelf units – approx. $1,000 - 2,000+/year.
Software: • Bundled with instrument. • Stand alone.	• Incorporated into cost of purchase. • $500 - $5000. Some have annual fees for version updates and technical support.
Sample throughput: • Manual placement on and off system (about 80 - 120 samples/hour). • Semi-automated (i.e., Bruker carousel holds 30 samples at a time, processing about 140 - 180 samples/hour).	• Requires labour to manually place and remove sample. • Requires labour to load and unload sample wheel.
Initial calibration model development: • Sample (otolith) collection costs. • Reference method costs (visual age estimation) to develop algorithm. • Species specific models. • Incorporation of spatial and temporal variability. • Chemometric analysis and model development (algorithm development). • Ongoing model maintenance.	• Species and location dependent The current project used 10 otoliths per age class for each species, season and region to develop calibration models and validations sets. However, calibration models approaching 1000 replicates are more common when developing a model where greater accuracy is required (Forrest *et al.* 2012). The grain industry, which analyses a high moisture sample, uses calibration models built over 5 to 7 years to incorporate sufficient between year variability for calibration models to be considered robust.

Discussion

The current project has made significant progress in applying NIRS to the science of ageing fish using otoliths. It should be noted that apart from the preliminary study by Wedding *et al*. (2014), the current project is the first dedicated research project to apply NIRS to fish ageing with structured input from several fisheries agencies involved in routine ageing of fish for management purposes.

Objective 1. Evaluation of NIRS to age fish

One of the project's objectives was to evaluate NIRS as a reliable, repeatable, cost-effective method of ageing fish. The project established that NIR spectra could be readily collected from fish otoliths and NIR calibration models developed that relate NIR spectra to fish age. However, predictive performance varied between species and geographic region, as well as between generic NIRS waveband models and geographic specific NIRS waveband models.

For Barramundi, geographic samples using a generic NIRS wavelength calibration model on 'fresh' otoliths had robust predictive performance in terms of R_v^2 (i.e., 0.85 to 0.90) and RMSEP (i.e., 6.6 and 7.5 months, see Table A3.1). Refined Barramundi models for fish ≤120 months had variable performance in terms of:

- the percentage of a sample correctly allocated to its age class (range 39% to 73%);
- the percentage of a sample whose predicted age was within 6 months of its observed age (range 57% to 73%);
- the percentage of a sample whose predicted age was within 12 months of its observed age (range 91% to 94%);
- between method levels of precision (i.e., IAPE range 3.4% to 6.9%);
- the absence of bias in the NIRS age predictions for either the Archer or Fitzroy fresh samples;
- a significant difference between the observed and NIRS predicted age-class distributions for the Fitzroy River samples (p=0.039), but not the Archer River samples (p=0.200).

Across most model configurations and areas, the overall NIRS model performance was better for Archer River Barramundi than for Fitzroy River Barramundi.

For Snapper, geographic samples using a generic NIRS wavelength calibration model on 'fresh' otoliths had robust predictive performance in terms of R_v^2 (i.e., 0.79 to 0.93) and RMSEP (i.e., 14.8 and 19.0 months, see Table A2.4). Refined Snapper models for fish <156 months had variable performance in terms of:

- the percentage of a sample correctly allocated to its age class (range 35% to 51%); the percentage of a sample whose predicted age was within 6 months of its observed age (range 32% to 52%);
- the percentage of a sample whose predicted age was within 12 months of its observed age (range 54% to 77%); and
- between method levels of precision (i.e., IAPE range 4.6% to 6.3%).

Age-bias plots indicated there are issues of age-bias for some but not all samples and there were no significant differences between the observed and NIRS predicted age-class distributions for any of the 'fresh' Snapper samples. In general, results suggested better NIRS model performance for Snapper from the Gulf St Vincent (SA) than from the Sunshine Coast (Qld).

In general, Snapper generic calibration models had better predictive performance than Barramundi generic calibration models, based on R^2, SDR, bias and slope. However, Snapper generic wavelength calibration models had a substantially higher predictive error than Barramundi generic calibration models. The substantial differences in NIRS predictive performance between Barramundi and Snapper indicates that some species are likely to have better NIRS predictive performance for fish age than others. The causes for differences in performance between species is unknown, but we suspect that it is a consequence of variability in otolith microchemistry included in the calibration sets as well as variable accuracy in the age estimation of the standard fish ageing method.

Efficacy of NIRS to predict fish age

It was difficult to evaluate the efficacy of NIRS to predict fish age because:

(i) the current study used a ‘proof of concept’ approach;

(ii) models were limited to generic wavebands selections and were not optimised for species by regions;

(iii) sample sizes were relatively small which limited extrapolation of results to full annual samples (and therefore comparison of metrics such as derived age-length keys or mortality rates from catch curve analyses); and

(iv) acceptable thresholds for fisheries performance measures such as percentage correct age-class allocation, degree of age-bias and levels of between method precision (i.e., IAPE) differ between different fisheries agencies.

In general, NIRS calibration models should meet the predictive performance requirements of the end user. The fisheries efficacy performance measures of the percentage of a sample whose predicted age was within 6 months of its observed age, the percentage of a sample whose predicted age was within 12 months of its observed age, the percentage correctly allocated to its observed age class, between method precision (i.e., IAPE), the presence (and tendency) of age-bias and comparison of age-class distributions were attempts to quantify how closely the NIRS predicted ages were to the observed ages. As discussed early (see Assumption and Limitations: Estimated age versus absolute age), the observed biological ages used as the reference method in developing the NIRS calibration models are likely to be under or overestimates by up to about three months, depending on when any particular individual was spawned compared to the nominal birth date Such errors will perpetuate through the NIRS calibration models and influence the NIRS predicted ages. However, this level of (in)accuracy is the reality of age estimation from otoliths of wild fish.

Fisheries performance measures for Barramundi were variable between locations (and sample years), with Archer River Barramundi having the best predictive performance. If based solely on results from the Archer River, it could be strongly argued that NIRS was an appropriate supplementary method to standard ageing methods. However, some of the other results suggest further research is warranted before NIRS can be recommended for widespread adoption in fish ageing. It was somewhat surprising that Fitzroy River Barramundi had lower predictive performance than Archer River Barramundi, despite Fitzroy River Barramundi having slightly better readability than Archer River Barramundi (i.e., 95% of the 2012 Fitzroy River Barramundi sample had the best readability score whilst 84% of the 2012 Archer River Barramundi sample had the best readability score), validation of annual increment formation (Stuart and McKillup 2002) and validation of increment counts from known age fish (Staunton-Smith *et al.* 2004). Therefore, it is possible that variability in fisheries performance measures for Barramundi are not the result of accuracy or precision issues in the visual estimates of fish age, but rather reflect variability in otolith microchemistry that probably accompanies a species that opportunistically uses freshwater, estuarine and marine habitats at varying stages throughout its life cycle. It may also be the consequence of variable post-mortem handling procedures and subsequent effects on otolith chemistry (see below for further discussion).

Snapper has variable certainty around its visual age estimates, with Snapper from the Gulf St Vincent reported to have greater certainty in age estimates (i.e., 34% excellent and 47% confident readability for the 2012 sample) than Snapper from the Sunshine Coast of Queensland (i.e., 54% very high and 46% medium readability for the 2012 sample) or Mid-West Coast of Western Australia (W. Sumpton pers. comm. 2014). Fisheries performance measures for Snapper were considerably lower than the results for Barramundi and were variable for Snapper from different locations (and sample years). Gulf St Vincent Snapper had the best predictive performance, with 77% of individuals predicted to within 12 months of their observed age. Age-bias in NIRS age predictions was more prevalent in Snapper than Barramundi and would need to be monitored, assessed for acceptability and adjusted for (possibly by data transformation) where necessary. The presence of age-bias and the values of IAPE (as a measure of between method levels of precision) recorded for Sunshine Coast Snapper would be considered unacceptable in an ongoing monitoring program such as that run by Fisheries Queensland (Stephen Wesche, pers. comm. 2014). Further research (see Recommendations) into applying NIRS to age fish may identify the causes of some of these issues and sufficiently resolve these problems such that NIRS has levels of efficacy that are acceptable to end-users within fisheries agencies. It would also be useful to have documented standards from fisheries agencies on what are acceptable levels of accuracy and precision for fish age estimation that would assist in setting the benchmarks that proxy methods such as NIRS need to achieve in order to gain widespread acceptability.

Objective 2. Effects of geographic and seasonal variation on NIRS

Geographic effects

Geographic effects were evident in the results. The predictive performance of calibration models was reduced when individual geographic location models were used to predict fish age of other geographic locations e.g., Archer River Barramundi calibration model used to predict Fitzroy River samples, R_v^2 = 0.51, RMSEP = 16.2 months, see Table A3.3; and Sunshine Coast Snapper calibration model used to predict Gulf St Vincent samples R_v^2 = 0.63 and RMSEP = 34.4 months (data not shown).

Geographic effects are not unusual in NIRS, where biological properties of agricultural produce often vary geographically and seasonally. In otoliths, differences in location specific NIRS calibration models probably reflect spatial differences in the (micro) chemistry of otoliths. In some cases, spatial variability could be addressed through applying a bias correction, although a better solution maybe to include all locations in the calibration set thereby incorporating the variability in microchemistry that is likely to occur in the prediction set.

Pooling fresh Snapper samples (i.e., Queensland and South Australia, Summer, Autumn, Winter and Spring) to produce a generic Snapper calibration model incorporated a wider range of otolith variability. This produced a more robust NIRS calibration model that had predictive performance comparable to geographic specific NIRS models (i.e., R_v^2 = 0.85; RMSEP = 17.9 months, see Table A3.4). Pooling all fresh Barramundi samples (i.e., Archer and Fitzroy Rivers, February, May and October 2012) produced a model with R_v^2 of 0.87 and RMSEP of 7.8 months. These results are comparable to the location specific models for Barramundi.

Calibration models for Gulf St Vincent Snapper had higher predictive performance than those for the Sunshine Coast, Queensland. This probably reflects the amount of spatial variability in the microchemistry of Snapper otoliths and/or the differences in the reference method of age estimation. The results suggest that Snapper otoliths sampled from the Sunshine Coast, Queensland maybe more variable in their microchemistry than those of the Gulf St Vincent, and may reflect the diversity of environments from which the samples were drawn (Gillanders 2002). Alternatively, it may be a consequence of: (i) inherent differences in the accuracy of the visual age assessment (i.e., reference method), with South Australian Snapper having better accuracy than Queensland Snapper; and/or (ii) post-mortem handling techniques. The South Australian samples of Snapper were sonicated in water prior to NIR spectra acquisition (Table A2.3), possibly leaving inert (age-dependent) compounds tightly bound to the calcium carbonate matrix of the otolith intact and uncluttered by variable water soluble components.

Pooling of samples across locations (and stocks) increased the predictive performance of NIRS models in some instances. However, stocks may also cover a wider geographic range than was assessed by the current project. For example, the southern Gulf of Carpentaria Barramundi stock Figure 55) covers a much wider geographic range than the Archer River Barramundi analysed by the current project. Potential geographic affects need to be identified and considered so that the appropriate spatial scale of samples can be included during development of NIRS calibration models and the potential spatial scale of unknown samples to be predicted.

Seasonal effects

There were no significant within year seasonal effects on the predictive performance of NIRS calibration models for Fitzroy River Barramundi (i.e., February and October 2012). Likewise for Snapper from the Gulf St Vincent, the predictive performance for models using all months was similar to models built for individual seasons. However, it is important to include within year seasonal variability in the calibration set of otoliths if this is likely to be a feature of the prediction set.

Objective 3. Effects of otolith storage time on NIRS

The repeat acquisition of NIR spectra from ‘fresh’ otoliths (up to 17 months post collection) suggest that otolith chemistry stabilises somewhere between six and 11 months following collection for Barramundi and about six months following collection for Snapper. The time difference in otolith stabilisation maybe the consequence of differences between species and spatial environmental conditions, as well as differences in the collection, processing and storage procedures of different fishery agencies.

Stabilisation of the ‘fresh’ otolith samples is possibly the consequence of amino acids within the otolith denaturing over time, or pH changing when fluids and ions are lost as moisture levels stabilise during storage. These issues could be potentially addressed by drying the ‘fresh’ otoliths in a drying oven at a predetermined temperature and for a set time period to ensure otoliths have stabilised prior to NIR spectra collection.

NIR spectra acquired from ‘historical’ Barramundi otoliths (stored for up to 105 months since collection) showed highly variable statistical results. There were no clear trends to indicate further stabilisation or degradation in otolith chemistry over storage time. This suggests that the main source of variation in the samples assessed was temporal differences (i.e., yearly) as each historic data set was obtained from a different year. Temporal and spatial effects have major impacts on the robustness of NIRS calibration models, but can be accommodated by ensuring the calibration sample set includes otoliths collected from different locations and in different years and months. The development of robust calibration models requires training sets that cover variables such as seasonal otolith growth, yearly (temporal) differences, fish size, diversified age structure and location variables, and measurement conditions (sample handling and presentation).

In comparison, NIR spectra acquired from ‘historic’ Snapper otoliths (stored for up to 85 months since collection) showed minimal temporal differences, suggesting similarity in otolith chemistry of samples collected in different years. One possible explanation for this consistency is the relatively stable environmental conditions of Snapper (i.e., salinity of offshore water), as compared to the variable environment of Barramundi (i.e., salinity of estuarine waters, as well as the use of freshwater habitats by a variable proportion of the Barramundi population).

Unfortunately, the samples of ‘fresh’ and ‘historic’ otoliths analysed by the current project were insufficient to answer the question of otolith degradation over storage times from 1 to >5 years, because of the confounding effect of temporal variability. Further work would be required to elucidate temporal variability from otolith degradation (see Recommendations).

Objective 4. Evaluate cost-effectiveness of ageing fish by NIR spectrometry

As previously identified, otoliths are generally more expensive to collect than to process for visual age assessment via thin sections. However, few agencies (except WA Fisheries, see Craine *et al*. 2009) have assessed whether their fish length and age monitoring programs are sufficient (i.e., collecting too much or not enough information) to inform further analyses based on fish age and length. Craine *et al.* (2009) compared numerous alternative (proxy) measures of age (i.e., otolith weight and total length) as an alternative to using the annuli method. Despite good relationships between otolith annuli method and various proxy measures, sample numbers for proxy methods required to achieve a Kolmogorov-Smirnov precision level of 0.04 were always considerably greater than sample numbers for visual age assessment via thin sections of otoliths. In most cases, Craine *et al.* (2009) found that proxy methods were not a cost effective alternative to standard fish ageing. The exception was for species where bulk samples could be collected cheaply e.g., cartons of pilchards purchased at the metropolitan fish market.

The example of the hypothetical running costs of supplementing standard ageing methods with NIRS suggested significant cost savings could be potentially achieved. The scale of the cost savings will depend on how variable NIRS with fish age is between years and the number of otoliths that need to be added to the calibration model as part of ‘model maintenance’.

It should be noted that developing and validating a NIRS calibration model for fish age based on otoliths is likely to require considerable time (i.e., greater than five years), effort and cost (see ‘NIRS spectroscopy costs’ and Table 21) before the model is sufficiently robust to only require ongoing ‘model maintenance’.

Much of the NIRS model development and validation could be outsourced to dedicated NIRS work-groups within or external to a fisheries agency, that may assist in minimising capital expenditure costs. Many fisheries agencies are part of large organisations servicing primary industries, that often have NIRS expertise albeit in other applications (e.g., food quality, malting standards of wheat and barley for beer, pasture nutrition etc.), that could be utilised cost-effectively, whilst retaining connection with the generation of fish age estimates. However, developing in-house NIRS capability may be desirable if large numbers of otoliths are processed annually or an agency wishes to retain control of all data estimation processes.

Other issues

Effects of NIRS of fish age on recruitment indices and YCS calculations

For some species, the age structure of a stock is used to calculate recruitment indices, sometimes referred to as Year-Class Strength (YCS), which are often based on the residuals from catch curve analyses of proportions-at-age. Examples relevant to the current project include Snapper in South Australia and Barramundi in Queensland and the Northern Territory (Staunton-Smith *et al.* 2004; Fowler *et al.* 2010; Halliday *et al.* 2014). In these calculations, it is important that age structures are an accurate estimate of the absolute age (and subsequent back-calculated year-class), because these indices of recruitment are often correlated to environmental drivers at annual time scales, such as temperature for Snapper in South Australia or river flow for Barramundi in northern Australia. It is unknown whether the predictive estimates of fish age based on NIR spectra of whole dry otoliths offers sufficient accuracy in estimating absolute age upon which to base analyses of year-class strength. NIRS fish age estimates are influenced by the accuracy of the standard visual ageing of the calibration sample, but will also be influenced by the relationship that we currently measure as a correlation between fish age and NIR spectra. Until issues of accuracy and how NIR spectra are related to fish age are better understood, it is recommended that application of NIRS estimate of fish age for calculation of recruitment indices be treated with some caution.

What does NIR spectroscopy measure in fish otoliths that is related to fish age?

The current study was designed as a ‘proof of concept study’ of NIRS to age fish, building on the preliminary results recorded for Saddletail Snapper (Wedding *et al.* 2014). Spectral information is directly related to the chemical structure of the otoliths and is assumed to be related to otolith age. However, what specific chemical component(s) in the otolith that the NIR spectra correlate to, are at this stage unknown.

Otoliths are calcified structures, with the deposition rate of aragonite, organic matter and trace elements varying with ontogenetic stage, and also as a consequence of environmental factors such as water temperature, salinity and the chemical composition of the water in which the fish swim (Campana 1999; Gillanders 2002). The variability in the performance of NIRS to predict fish age (between species and regions) caused most of the fisheries collaborators in the current project to want to know “What does NIRS measure in fish otoliths that is related to fish age?” with the subsequent question of “Is the correlative relationship between the NIR spectra of otoliths and fish age dependent on time or growth (rate) or both?”. Further research into what NIRS measures in fish otoliths would assist in refining methods for estimating fish age via NIRS and would improve the confidence of end users in the applicability of the calibration models.

However, the question of what NIRS measures in relation to the quality or compositional attribute of interest is never answered or fully confirmed in many industries that routinely use NIRS technology. It is not a critical requirement of the NIRS technique as to whether there is a direct or secondary correlation between the chemical composition (of the reference material) and the property of interest. However, it is important that the predictive performance of the NIRS calibration models meet the requirements of the end user.

Post-mortem handling effects on otolith chemistry and implications for NIR spectra

It became apparent during the current project that post-mortem handling procedures applied to fish otoliths varied greatly as a consequence of: (i) collection method; and (ii) agency protocols (Table A2.3). Effects of post-mortem handling on fish otoliths has been studied (Gauldie *et al.* 1998; Milton and Chenery 1998; Proctor and Thresher 1998; Rooker *et al.* 2001; Hedges *et al.* 2004), but the literature is insufficient to determine what the effects are likely to be on NIR spectra – except that handling procedures are probably one cause of variability. The consequences of common post-mortem handling procedures (i.e., freezing, duration of in-vitro storage before otolith extraction, holding temperature, cleaning procedures) on the (macro and micro) chemistry of otoliths needs to be understood (and standardised) because NIRS is a secondary method of determination and consistency in (calibration and validation) sample handling is likely to produce more accurate predictions. Time differences in the stabilisation of NIR calibration models from the ‘fresh’ otolith samples observed in the current project maybe the consequence of differences between species and/or collection location, as well as differences in post-mortem handling procedures.

Potential applicability of NIRS

NIRS has also been applied to the age estimation of Chondrichthyan vertebrae, dorsal fin spines and skin (Rigby *et al.* 2014; Rigby *et al.* In review (a)). NIRS may also be applicable to calcareous tissues such as spines. If so, NIRS would be highly useful in providing age estimates of species where otolith removal is not generally possible but fin clips are (e.g., live coral trout or grey mackerel). NIRS would then also offer the opportunity to increase the proportion of harvested population sampled for age as fin clips and/or spine samples are more readily obtainable than otolith samples. Thus, the NIRS method maybe applicable to calcareous tissues of live as well as dead specimens, with potential future developments for non-lethal in-field ageing of live fish using hand held NIR units.

Conclusion

The current project has delivered on all four project objectives by innovatively merging the disparate science disciplines of NIR spectroscopy and fisheries to evaluate the ability of NIRS to age fish, specifically for Barramundi from northern Australia and Snapper from southern Australia.

Key findings for each of the project objectives are provided below.

Objective 1. Evaluate Near Infrared Spectrometry (NIRS) as a reliable, repeatable, cost-effective method of ageing fish

Results of the current project demonstrate that NIRS can be used as a predictive tool to estimate fish age based on NIR spectra from otoliths and has potential to supplement standard ageing methods, pending further research. The predictive performance of NIRS calibration models will vary between fish species, with the accuracy of the reference method playing a significant role in this variability. The applicability of NIRS to age fish species will depend on the species and location of interest, as well as the level of efficacy in predicted ages acceptable to a fisheries agency. NIRS may be particularly useful for fish species that are difficult to age, such as small pelagic species (e.g. pilchards). NIRS may also be applicable to other calcareous features of fish such as spines and fin rays, and with further development may potentially offer non-lethal in-field ageing of fish using hand held NIRS units.

The advantages of estimating fish age using NIRS include:

- Reduced costs compared to standard ageing methods;
- Non-destructive assessment of whole otoliths;
- Assessment can be rapid;
- Can be used to estimate age not only on whole otoliths, but also damaged or chipped otoliths;
- Minimal sample preparation;
- Multi-analytic (i.e., the ability to assess multiple attributes from a single NIR spectrum); and
- Objective and of high precision.

The disadvantage of estimating fish age using NIRS include:

- Unknown errors in the visual age assessment of samples plus any NIRS associated errors are propagated through to the predicted age estimates;
- Potential for age-bias in some species or locations, particular underestimating the age of older fish, but may be addressed through data transformation;
- Although NIRS prediction is rapid, calibration model development is dependent on the collection and standard ageing of otoliths, which is often not rapid;
- Development of robust calibration models is time-consuming and laborious;
- Current uncertainty if the correlation between NIR spectra and fish age is independent of fish growth, but could be addressed through further research;
- Current uncertainty in what NIRS measures in fish otoliths that is related to age, resulting in lowered confidence in applicability of NIRS to age fish by most fisheries scientists, but could be addressed through further research.

Objective 2. Determine the effect of geographic location (including latitude) distribution on NIRS algorithm stability

Geographic effects on the stability of NIRS calibration models were evident, being most prominent when models built for one location were used to predict fish age from another location. Geographic effects are not unusual in NIRS applications to agricultural produce (e.g., grains and fruits), where the biological properties of interest often vary geographically and seasonally. However, geographic effects could be accounted for in NIRS calibration models by including samples from all locations of interest and producing a robust model that had predictive performance (R^2 and RMSEP) comparable to geographically specific models. Season (within year) had little effect on the predictive performance of NIRS models for Barramundi or Snapper, although seasonality should always be considered as a contributor to variability in NIRS. We suspect that variability in results is probably a consequence of variable otolith microchemistry and how this differs between species, locations and year of collection, as well as differences in the collection, processing and storage procedures of otoliths (i.e., post-mortem handling).

Objective 3. Determine the effect (if any) of otolith storage time (years/months) on NIRS estimates of age

The effects of storage on NIRS estimates of fish age were assessed by repeatedly acquiring NIR spectra from ‘fresh’ otoliths up to 17 months after collection. Results suggest that otolith chemistry stabilises somewhere between six and 11 months following collection for Barramundi and at about six months following collection for Snapper. The time difference in stabilisation may be the consequence of differences between species and their associated environmental conditions (e.g., fresh/estuarine habitats compared to oceanic habitats), as well as differences in post-mortem handling of otoliths. Samples of ‘fresh’ and ‘historic’ otoliths analysed by the current project were insufficient to determine if otoliths degrade over storage times from one to more than five years, because of the confounding effect of between-year variability. Further work would be required to elucidate temporal variability from otolith degradation.

Objective 4. Evaluate the cost-effectiveness of ageing fish by NIRS vs. standard otolith ageing, and develop optimised fish sampling regimes with respect to ‘cost’ (defined in terms of labour, lab time, field costs, etc.)

It was difficult to evaluate the cost-effectiveness of ageing fish by NIRS, because the models developed in the current study were intended as ‘proof of concept’, were generally focused on generic waveband selection rather than optimising the model for regions or seasons, and it is difficult to predict what level of NIRS calibration model maintenance might be applicable. Therefore, it was difficult to “develop optimised fish sampling regimes with respect to “cost” as per objective 4 in the original project proposal. In addition, fisheries end-users commented that they would not really consider cost effectiveness until further work resolved some of the issues identified, such as the benefits of split age models, transformation of age data prior to analysis to address issues of age-bias and how age-length variability could be incorporated into NIRS model development and subsequent NIRS model maintenance. Case studies of hypothetical running costs suggest significant cost savings could be achieved if NIRS is used to supplement standard fish ageing methods. However, there are considerable start-up costs over a number of years (i.e., three to five plus years) to develop and validate NIRS calibration models. It may be more feasible for fisheries agencies to outsource to dedicated NIR spectroscopy groups than to develop their own capability. Many fisheries agencies are part of large organisations that have expertise in NIRS servicing primary industries that could be tapped into, thus minimising capital expenditure costs, whilst retaining connection with the generation of fish age estimates.

Applying NIR spectra to age fish is a novel and interesting development in the field of fish ageing, potentially offering a new way of looking at some of the complicated and protracted fisheries problems. For example, in the current project, principal component analysis of NIR spectra was able to discern Snapper caught in the Northern Gulf St Vincent from those caught in the Southern Gulf St Vincent. This suggests that provided appropriate calibration models can be developed, that NIRS may have applicability in the spatial discrimination of fish, answering questions such as stock structure, source of natal habitat and stocked versus wild fish.

Implications

Fish age can be estimated through the application of NIR spectroscopy to fish otoliths and potentially offers significant cost savings. However, further work is required to elucidate the causes of variability in the relationship between fish age and NIR spectra and minimise error in age predictions to a level acceptable to potential end users (e.g., State and Commonwealth fisheries agencies). The potential applicability of NIRS was recognised by end-user stakeholders in Queensland and the Northern Territory who are considering further research: (i) pilot application of NIRS 2012 Archer River Barramundi calibration models to the 2012 southern Gulf of Carpentaria barramundi length and age data; and (ii) stock discrimination of tropical coastal reef fish in the Northern Territory.

Standard ageing of fish otoliths via thin section is widely accepted as the current best estimate of observed fish age in Australia and overseas, despite its labour intensiveness, cost and levels of accuracy and precision. NIRS potentially offers a cost-effective alternative. However, each fisheries agency considering NIRS would need to assess whether ageing fish using NIRS offers sufficient cost-effectiveness, accuracy and precision to meet their data requirements.

Recommendations

Applying NIRS to fish otoliths is an innovative science that is worthy of further research, but there are several key issues that need to be addressed before NIRS can be recommended for supplementing standard fish ageing techniques in ongoing fisheries monitoring programs.

The key questions asked by fisheries collaborators in the current project were:

- What does NIRS measure in fish otoliths that is related to fish age? and
- Is the correlative relationship between NIRS and fish age dependent on time or growth (rate) or both?

Research that answers these questions would identify sources of error that could be minimised, identify the potential influence of growth rates on the relationship between fish age and NIRS, and guide post-mortem handling of otoliths for NIRS. This knowledge would lead to greater confidence by fisheries stakeholders in NIRS estimates of fish age and the suitability of NIRS to supplement standard fish ageing methods. Greater understanding of what NIRS can measure in fish otoliths would potentially broaden the applicability of NIRS as a tool in fisheries science e.g., to assist in spatial discrimination.

Further development

Pilot application of NIRS to the 2012 Southern Gulf of Carpentaria Barramundi age and length data set

Fisheries end-users in Queensland were keen to have the results for the 2012 Archer River Barramundi applied to the complete 2012 set of age and length data for the southern Gulf of Carpentaria stock of Barramundi, where 605 otoliths were collected and aged and 2503 Barramundi were measured for length. This ‘pilot application’ would involve scanning and predicting the age of all otoliths collected from the southern Gulf of Carpentaria based on the calibration models built in the current project. Where necessary, additional individuals may be incorporated into the current NIRS calibration model to incorporate additional variability (between NIRS and fish age) within the southern Gulf of Carpentaria stock. Age predictions based on NIRS could then be used to calculate the annual age-length key and upscaled to measured length data to generate the estimated age structure of this stock.

A full pilot application of NIRS to age fish of a whole stock would assist in identifying how this technology could be applied to ongoing monitoring programs. It could determine if some of the issues raised in the current project (e.g., split age models; stratification of calibration sample sets by length) are problematic (or not). Results could be compared between the standard ageing method and NIRS supplemented method based on the simplistic fisheries performance measures used in the current project, through to more complicated end-point results such as mortality estimates and estimated age structure of the stock. This would further inform potential pathways for adoption of NIRS into standard fish ageing programs.

Applicability of NIRS to spatial/stock discrimination

NIRS is widely used to determine the chemical composition of materials in a wide range of science disciplines. One of the most interesting results of the current project was the use of NIR spectra to discriminate between Snapper caught in the Northern and Southern Gulf St Vincent. The identification of macro constituents is the most common use of NIRS technology, as the sensitivity limit is about 0.1% for most constituents. However, NIRS has been used to detect constituents at very low levels (i.e., ppm and ppb). Often these low levels are not detected directly, but detected through secondary correlations. Therefore, provided a calibration model can be cost-effectively constructed, NIRS may have potential in the field of spatial discrimination of fish, such as stock discrimination and source/natal-habitat identification. Otolith microchemistry is expensive and often deals with limited sample numbers. NIRS offers the opportunity of greatly expanding the sample sizes analysed for otolith microchemistry (albeit by a secondary method), and thus potentially expand our ability to answer research questions (e.g., contribution of stocked fish to wild populations). Interest in this aspect of NIRS is demonstrated by the desire of NTDPI to explore the use of NIRS for stock structure of tropical coastal reef fish as part of FRDC project 2013/017 (T. Saunders pers. obs.)

Determining what NIRS measures in fish otoliths, how this related to fish age and implications for optimal post-mortem handling of otoliths for NIRS

Consensus amongst fisheries stakeholders consulted by the current project identified that determining what NIRS measures in fish otoliths and how it is related to fish age was a priority area for further research. This would involve breaking fish otoliths into their macro and micro constituents and determining the NIR spectral signature of each constituent (i.e., the 'wet chemistry'). It would also involve determining how quantities of these constituents change with fish age and growth rate (i.e., variation in what NIRS measures with age) to inform appropriate data transformation prior to NIRS calibration model development. This information could also be used to determine the optimum (or standardised) post-mortem handling procedures of otoliths for NIR spectra collection. Known age fish (such as aquaculture Barramundi from multiple locations to account for spatial variability in water chemistry) or a surrogate species (e.g. *Gambusia*) that is short-lived with fast early growth rates that plateau within the time frame of an academic study would assist in determining how NIRS is related to fish age.

Stability of NIR spectra from fish otoliths over time

One of the objectives of the current project was to determine the effect of storage time on NIRS estimates of fish age. We were able to partly answer this question; identifying that fresh otoliths did indeed stabilise for NIR spectra over six to 12 months (species and collection area dependent). However, it was uncertain as to whether historic otoliths had degraded over longer storage times because effects due to storage time and collection year were confounded. Further repeat acquisition of NIR spectra would contribute to answering how and whether otoliths degrade over the long term, thus impacting on NIRS calibration models. The current project repeatedly acquired NIR spectra for up to 17 months from one species from one geographic location. Extending this data collection, so that spectra were acquired at regular intervals up to five years after collection, would greatly assist in determining NIR spectral differences due to otolith degradation versus temporal variability.

Application of NIRS to 'difficult to age' species such as small pelagic species

NIRS may be useful in supplementing the age estimates of 'difficult to age' species, such as sardines, providing an alternative method (to otolith weight or fish length based analyses) to estimate of age of individuals with poor otolith readability.

Improving the performance of NIRS to age fish through more specific wavelength models and alternate data transformation

The majority of models developed in this "proof of concept" study were based on a generic set of wavelengths which were independent of temporal or regional effects. Areas in the spectrum where maximum differences occur in response to changes in concentration of substances of interest can indicate that these spectral regions or wavelengths will likely be used in an accurate analysis. Spectral region selection can significantly improve the performance of full spectrum calibration techniques and may be one of the most critical aspects of NIR analysis. Specific spectral regions are selected where co-linearity is not so important, generating more stable models with superior interpretability. In practice, this requires the identification of a subset of the full spectrum that will produce the lowest prediction error. By identifying specific wavelengths, the predictive performance of individual calibration models can be improved and model robustness enhanced.

All NIR spectra discussed in the current project were mathematically transformed prior to PLS regression analysis. Transforming data prior to analysis assists with removing defects observed in the NIR spectra (e.g., noise, base line drift). A 25-point Savitsky-Golay (SG) spectral smooth (2nd order polynomial) followed by a first derivative transformation (25-point SG smooth and 2nd order polynomial) were applied to all spectra. Improved model performance could be obtained by investigating alternative data transformations and selecting the most appropriate for each individual location by season data set.

Extension and Adoption

The project idea and initial results from the Saddletail Snapper work were presented at the 2011 workshop on fish ageing associated with the annual conference of the Australian Society for Fish Biology. The workshop was attended by fish ageing representatives from all Australian state and territory fisheries agencies, as well as the Commonwealth.

The 2011 workshop resulted in four state fisheries agencies (Queensland, Northern Territory, Western Australia, and South Australia) having co-investigators on the project to assist in project development, interpretation of results and dissemination of the potential of NIRS to age fish within their respective agencies.

Results from the current project were presented (by Brett Wedding) and discussed at the 2013 workshop on 'Routine fish ageing' associated with the 2013 joint meeting of the Australian Society for Fish Biology and the New Zealand Marine Sciences Society held in Hamilton New Zealand on the 18th of August 2013. This workshop was attended by fisheries scientists and managers from most states of Australia, as well as key fisheries scientists from New Zealand.

A project workshop of participating agencies was held in Brisbane on 29 May 2014, (attendees listed in Appendix 3), with an invitation extended to Fish Ageing Services. Discussion included principles of NIRS, project results, methods to compare the cost-efficiencies of the alternate ageing methods, and further R&D.

Guidelines to developing protocols for incorporating NIRS in fish ageing programs is an output of the project (see Appendix 4) and will be made available to all fisheries agencies in Australia, as well as any other interested parties.

The project team have intentions of publishing the results in a scientific journal to communicate the potential for NIRS to age fish to the wider scientific community.

Project coverage

- **Media Coverage**

A media statement on the project was released by DAF on 15 November 2012. The project received additional media coverage in the Cairns Post on 1 June 2012.

- ***Conferences***

Brett Wedding (Project Co-Investigator) presented project results at the workshop on "Routine fish ageing using otoliths for fishery monitoring and stock assessment" held in Hamilton New Zealand on 18 August 2013. This workshop was held one day prior to the joint conference of the Australian Society for Fish Biology and the New Zealand Marine Sciences Society. Many of the workshop attendees were interested in the possibility of supplementing traditional fish ageing methods (i.e., sectioning) with rapid technology, such as NIRS calibration models. The presentation was very well received and generated great interest in progressing future collaborative projects to enable industry productivity gain.

Carole Wright (Project Staff) presented project results as a poster at the Australasian Applied Statistics Conference 2012 (GenStat & ASReml) 3-7 December 2012 in Queenstown, New Zealand, entitled "Using NIRS to estimate the age of tropical fish from otoliths", authors Carole Wright, Brett Wedding, Steve Grauf, Julie Robins, Michelle Sellin and Sue Poole.

Glossary

Statistical terms:

The standard definition of the ***coefficient of determination*** **(R^2)** is that it measures the extent to which the fitted straight line relationship explains the variability in the *y*-values. The R^2 calculated by the Unscrambler chemometric software package is the proportion of the total variance accounted for by the explained variance for the given number of latent variables. In both situations, R^2 has values between 0 and 1.The R^2 for the prediction samples is adjusted to account for the number of latent variables in the calibration model and is not equivalent to the square of the correlation coefficient. A low R^2 between NIR spectra and reference data indicates that the NIR analysis has not been successful, and it is possible that NIRS may not be applicable to that particular analysis. A coefficient of determination for a validation set of at least 0.65 is required to be able to sort agricultural commodities into at least two grades (i.e., above and below an acceptable level) with approximately 80% accuracy (Guthrie *et al.* 1998).

The ***root mean square error of prediction*** **(RMSEP)** is used to indicate the predictive performance of the calibration model. The RMSEP is an estimation of the variation between the reference value and NIRS predicted values of the validation or cross validation set i.e., it measures how accurately the predicted age compares to the actual age. It is usually calculated in the units of the reference material – which in the current project was biological age in months. An RMSE of 0 indicates a perfect fit between the predicted and reference age across the sample set. Increased RMSE indicates reduced fit between the predicted and reference age data. The RMSEP incorporates the error from the reference method (visual age assessment) and the NIR technique

For ***root mean square error of cross validation*** **(RMSECV)**, a leave-one-sample-out cross validation is performed. The NIR spectrum of one sample in the calibration set is excluded and a partial least square (PLS) model is built with the remaining spectra. The left out sample is predicted with this model and the procedure is repeated in turn by leaving out each of the samples of the calibration set. Cross validation of a calibration model makes it possible to select the optimum number of latent variables, that is, the number giving the minimum prediction error for the calibration set.

The ***root mean standard error of calibration*** **(RMSEC)** is an estimation of the variation between the reference value and NIRS predicted values of the calibration set, i.e., how well the calibration model fits the calibration set.

The ***bias*** corresponds to the average difference between reference values and NIRS predicted values. If there is no such difference, the bias will be zero.

The $\mathbf{R_c^2}$ ***(calibration model coefficient of determination)*** is a function of the standard deviation (SD) and RMSECV. The $\mathbf{R_v^2}$ ***(validation coefficient of determination)*** is a function of SD and RMSEP (when bias = 0). The R^2 can often be improved by increasing the SD of the calibration set. Therefore, an evaluation of a model using the R^2 statistic should be considered in conjunction with knowledge of the SD, which should be similar to that of sample set to be predicted.

The lowest RMSE and highest R^2 will be the optimum combination of wavelength and pre-processing treatment. Ideally, the RMSEP and the bias should both be equal to zero with a small difference between RMSECV and RMSEP. Thus, a high R^2 with a low RMSEP and bias means that the NIR results are accurate over the anticipated range, and likely to remain so, provided the statistics were based on a sufficient number of observations.

In general, calibration model error may easily double when a calibration model is applied to a spectral data set of a different season, year or location. This lack of robustness often translates into bias. Prediction bias for new sample sets can be corrected by model updating or direct bias adjustment. A ***bias adjustment*** is appropriate when all predicted values consistently read higher (or lower) than the measured (reference) values. A large positive (or negative) bias can be corrected by adding (or subtracting) the **absolute** bias value to all reference values and re-running the prediction model.

Latent variables: As more latent variables (factors or terms) are included in the calibration model, the model begins to fit the random errors embedded in the spectra and concentrations. Therefore, the RMSEC will always decrease as more factors are added and may end up much lower than the laboratory error of the reference method. When extra factors that mostly describe random errors are included in the calibration model, these factors will not fit the errors in the future samples and the RMSECV and RMSEP may increase. The best policy is to keep the number of wavelength terms as low as practicable. A simple rule of thumb for NIR chemometrics is to use a minimum of 5 to 15 samples for each regression and data treatment constant and for any parameter of the data treatment (such as wavelength) that is allowed to vary.

Wavelength selection: Areas in the spectrum where maximum differences occur in response to changes in concentration of substances of interest can indicate that these spectral regions or wavelengths will likely be used in an accurate analysis. Spectral region selection can significantly improve the performance over full-spectrum calibration techniques and maybe one of the most critical aspects of NIR analysis. Specific spectral regions are selected where co-linearity is not so important, generating more stable models with superior interpretability. In practice, this requires the identification of a subset of the complete data that will produce the lowest prediction error.

The ***standard deviation ratio*** **(SDR)** is the ratio of the standard deviation (SD) of the population divided by the RMSEP or RMSECV. The SDR statistic enables comparison of model performance across populations with different distributions and thus SD. In theory (for large normally distributed data sets), SDR is equal to and indicates more directly than either R^2 or RMSEP separately can, the relative predictive performance of a model; the higher the value, the greater the power. McGlone and Kawano (1998) suggest that an SDR of >3 is adequate to support sorting/grading into three classes, while Golic and Walsh (2006) suggest that an SDR of 2.5 allows sorting into two grades.

Project materials developed

Guidelines for the application of Near Infrared (NIR) Spectroscopy for predicting the age of whole dry otoliths

The project has developed “Guidelines for the Application of Near Infrared (NIR) Spectroscopy for Predicting the Age of Whole Dry Otoliths” (see Appendix 4) that provides a brief set of guidelines to assist potential users with the basis of developing their own protocols and calibration models for applying NIRS to predict the age of fish based on otoliths. It includes background information, NIR equipment, otolith sample presentation, NIR spectra acquisition, calibration model development, chemometric analysis and model evaluation.

Appendix 1: Project Staff

(in alphabetical order)

Department of Agriculture and Fisheries

Mr Steve Grauf, Senior Technician, Crop and Food Science	NIR expertise
Dr Julie Robins, Senior Fisheries Biologist, Fisheries & Aquaculture	fisheries expertise
Ms Michelle Sellin, Fisheries Technician, Fisheries & Aquaculture	fisheries expertise
Mr Brett Wedding, Principal Scientist, Crop & Food Science	NIR expertise
Dr Carole Wright, Senior Biometrician, Horticulture & Forestry Science	NIR expertise

Department of Primary Industries and Fisheries

Dr Thor Saunders, Principal Research Scientist	fisheries expertise

South Australian Research and Development Institute – Aquatic Sciences

Dr Anthony Fowler, Sub Program Leader, Marine Scalefish Research Group	fisheries expertise

Department of Fisheries, Western Australian

Dr Stephen Newman, Principal Research Scientist	fisheries expertise

Project Review Workshop (May 2014) participants

All project staff, plus:

Ms Sue Helmke, DAF
Dr Sue Poole, DAF
Dr Wayne Sumpton, DAF
Dr Jonathan Staunton Smith, DAF
Dr Stephen Wesche, DAF
Ms Olivia Whybird, DAF

Appendix 2: Otolith Collection and Processing Protocols

Table A2.1 Age-allocation matrix for Queensland Barramundi to convert sample date, visually assessed increment count and edge interpretation to age-class based on a nominal birth date of 1st January. IC = increment count

Edge Interpretation	Collection Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
New	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	IC-1	
Intermediate	IC	IC	IC	IC	IC	IC	IC	IC	IC-1	IC-1	IC-1	
Wide	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	

Table A2.2 Age-allocation matrix for Queensland Snapper to convert sample date, visually assessed increment count and edge interpretation to age-class based on a nominal birth date of 1st July. IC = increment count

Edge Interpretation	Collection Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
New	IC	IC	IC	IC	IC-1	IC-1	IC	IC	IC	IC	IC	IC
Intermediate	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC
Wide	IC	IC	IC	IC	IC	IC	IC+1	IC+1	IC+1	IC	IC	IC

Table A2.3 Post-mortem handling and otolith sample preparation of Barramundi and Snapper otoliths supplied to the current project

Species	Geographic Location	Sample Preparation Method
Barramundi	Archer River estuary, Queensland	• Otoliths are collected on board a commercial vessel over a period of about five days. • Otoliths are removed, blotted dry with tissue paper and placed into a plastic vial. • Otoliths are not washed in any solution. • The vials are left uncapped for at least two days with some tissue paper placed in them. • Otoliths collected on the last few days are placed in the fume hood for roughly two days to ensure they are dry. • Vials and otoliths are then checked before capping to make sure they are dry and for any discrepancies.
	Fitzroy River estuary, Queensland	• Washed in freshwater, blotted dry on paper towel and placed into plastic vials. • 40 of the 206 February 2012 otolith samples were cleaned again before sending for NIRS assessment. These samples were re-cleaned due to the presence of a black residue on the otolith. • These samples were brushed lightly with a new toothbrush, rinsed under running tap water, dried with tissues and placed into plastic vials
	Daly River estuary, Northern Territory	• Washed in fresh (or salt?) water, blotted dry on paper towel and placed into plastic vials.
Snapper	Sunshine Coast Offshore, Queensland	• Washed in freshwater, blotted dry on paper towel and placed into plastic vials.
	Gulf St Vincent, South Australia	• The fish are sampled at the SAFCOL fish market. • Typically otoliths are removed and place into plastic bags at the markets. • Back in the laboratory the otoliths are cleaned in water, wiped on paper towel and then stored in small plastic bags. • For this project otoliths were cleaned in an ultrasonic cleaner using distilled water, and placed in plastic vials for transport.
	Mid-West Coast, Western Australia	• For this project, otoliths were collected and stored in paper envelops and provided to a Queensland agency to pass on for NIRS assessment. • 90 of the 201 otoliths supplied were cleaned by the Queensland agency by lightly brushing with a (new) toothbrush and cold tap water. • 90 of the 201 otoliths were then briefly rinsed in running tap water (never allowed to soak). Blotted dry with lab tissue paper and placed into plastic vials ready for transport. • The otoliths that were cleaned had a black residue on them, usually in the sulcus. Only the area of the otoliths with the black residue was brushed with the toothbrush. • Otoliths were then placed into plastic vials ready for transport.

Appendix 3: Performance Results of NIRS Models

Table A3.1 Partial least squares statistics for biological age of ‘fresh’ Barramundi otoliths of all ages using a generic NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (months)	LV	R^2 (bias corrected)	RMSE (months) (bias corrected)	Bias (months)	SDR	r	Slope
Archer River, May 2012		94	28 – 148 (22.2)	3	0.88	7.6	-0.08	2.91	0.94	0.88
	Archer River, May 2012	90	28 – 112 (17.1)	3	0.85	6.6	2.19	2.61	0.94	1.04
	Fitzroy River, Feb 2012	204	25 – 145 (23.2)	3	0.63 (0.83)	14.1 (9.5)	-10.43	1.65 2.46	0.92	0.73
	Fitzroy River, Oct 2012	199	32 – 141 (22.3)	3	0.74 (0.86)	11.2 (8.3)	-7.59	1.98 2.69	0.93	0.84
Fitzroy River, Feb 2012		104	25 – 145 (25.3)	4	0.88	9.0	-0.20	2.82	0.93	0.88
	Fitzroy River, Feb 2012	100	25 – 109 (20.9)	4	0.87	7.5	0.07	2.80	0.94	0.99
	Fitzroy River, Oct 2012	199	32 – 141 (22.3)	4	0.86 (0.91)	8.3 (6.8)	4.84	2.68 3.30	0.96	0.98
	Archer River, May 2012	184	28 – 148 (19.9)	4	Not reportable (0.80)	27.1 (8.9)	25.58	0.73 2.24	0.93	1.09
Fitzroy River, Oct 2012		100	32 – 141 (22.7)	3	0.91	7.0	-0.04	3.26	0.95	0.90
	Fitzroy River, Oct 2012	99	32 – 141 (22.0)	3	0.90	6.8	1.17	3.26	0.95	0.96
	Fitzroy River, Feb 2012	204 (1)	25 – 145 (23.2)	3	0.85	8.9	-3.92	2.60	0.94	0.89
	Archer River, Feb 2012	184	28 – 148 (19.9)	3	0.56 (0.85)	13.2 (7.7)	10.72	1.51 2.59	0.93	0.99
Archer & Fitzroy, May & Feb 2012		198	25 – 148 (25.2)	2	0.86	9.6	-0.04	2.62	0.86	0.85
	Archer & Fitzroy, May & Feb 2012	190 (2)	25 – 112 (20.9)	2	0.86	7.6	1.76	2.76	0.94	1.01
Archer & Fitzroy, May, Feb & Oct 2012		298	25 – 148 (24.5)	2	0.86	9.2	-0.02	2.66	0.86	0.86
	Archer & Fitzroy, May, Feb & Oct 2012	289	25 – 141 (21.3)	2	0.87	7.8	1.20	2.74	0.94	0.96

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.2 Partial least squares statistics for biological age of 'fresh' Barramundi otoliths ≤120 months using a generic NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (months)	LV	R^2 (bias corrected)	RMSE (months) (bias corrected)	Bias (months)	SDR	r	Slope
Archer River, May 2012		94 (3)	28 – 112 (17.8)	2	0.86	6.7	-0.01	2.68	0.93	0.86
	Archer River, May 2012	90	28 – 112 (17.1)	2	0.88	5.9	1.07	2.90	0.94	0.87
	Fitzroy River, Feb 2012	204 (2)	25 – 109 (21.5)	2	0.27 (0.73)	18.3 (11.1)	-14.57	1.17 1.94	0.86	0.70
	Fitzroy River, Oct 2012	196 (3)	32 – 117 (19.8)	2	0.30 (0.87)	16.5 (7.1)	-14.93	1.20 2.80	0.94	0.78
Fitzroy River, Feb 2012		104 (2)	25 – 109 (22.2)	3	0.87	7.9	-0.07	2.81	0.93	0.88
	Fitzroy River, Feb 2012	100	25 – 109 (20.9)	3	0.86	7.7	-0.30	2.73	0.93	0.95
	Fitzroy River, Oct 2012	196	32 – 117 (19.8)	3	0.90	6.3	0.90	2.63	0.95	0.94
	Archer River, May 2012	184 (3)	28 – 112 (17.4)	3	Not reportable (0.81)	26.4 (7.6)	25.22	0.66 2.29	0.92	1.01
Fitzroy River, Oct 2012		98	32 – 105 (19.5)	4	0.92	6.0	-0.02	3.43	0.96	0.92
	Fitzroy River, Oct 2012	98	32 – 117 (20.1)	4	0.90	6.2	0.72	3.22	0.95	0.94
	Fitzroy River, Feb 2012	204	25 – 109 (21.5)	4	0.88	7.5	1.28	2.86	0.94	0.90
	Archer River, Feb 2012	184 (3)	28 – 112 (17.4)	4	Not reportable (0.80)	24.5 (7.8)	23.25	0.71 2.23	0.91	0.97
Archer & Fitzroy, May & Feb 2012		193	25 – 112 (21.7)	3	0.80	9.8	0.03	2.21	0.89	0.80
	Archer & Fitzroy, May & Feb 2012	190	25 – 112 (20.9)	3	0.82	8.8	0.42	2.38	0.91	0.85
Archer & Fitzroy, May, Feb & Oct 2012		292	25 – 117 (21.4)	3	0.81	9.2	-0.01	2.29	0.90	0.81
	Archer & Fitzroy, May, Feb & Oct 2012	287	25 – 112 (20.4)	3	0.82	8.7	0.30	2.35	0.91	0.81

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.3 Partial least squares statistics for biological age of 'fresh' Barramundi otoliths of all ages using a geographic location sample specific NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (month)s	LV	R^2 (bias corrected)	RMSE (months) (bias corrected)	Bias (months)	SDR	r	Slope
Archer River, May 2012		94	28 – 148 (22.2)	4	0.89	6.8	-0.12	3.27	0.94	0.90
	Archer River, May 2012	90	28 – 112 (17.1)	4	0.85	6.6	2.19	2.60	0.94	0.99
	Fitzroy River, Feb 2012	204	25 – 145 (23.2)	4	0.51 (0.83)	16.2 (9.7)	-13.07	1.43 2.41	0.91	0.75
	Fitzroy River, Oct 2012	199	32 – 141 (22.3)	4	0.78 (0.86)	10.3 (8.3)	-6.19	2.15 2.69	0.93	0.83
Fitzroy River, Feb 2012		104	25 – 145 (25.3)	3	0.88	8.8	-0.14	2.87	0.94	0.88
	Fitzroy River, Feb 2012	100	25 – 109 (20.9)	3	0.87	7.8	0.99	2.72	0.94	0.96
	Fitzroy River, Oct 2012	199	32 – 141 (22.3)	3	0.87 (0.91)	8.1 (6.6)	4.74	2.74 3.37	0.96	0.95
	Archer River, May 2012	184	28 – 148 (19.9)	3	Not reportable (0.83)	21.1 (8.2)	19.41	0.94 2.42	0.93	1.01
Fitzroy River, Oct 2012		100 (3)	32 – 141 (22.7)	2	0.92	5.5	-0.07	4.12	0.96	0.92
	Fitzroy River, Oct 2012	99 (1)	32 – 141 (22.0)	2	0.91	6.1	0.20	3.62	0.95	0.95
	Fitzroy River, Feb 2012	204 (1)	25 – 145 (23.2)	2	0.87	8.4	-1.73	2.75	0.93	0.85
	Archer River, May 2012	184	28 – 148 (19.9)	2	0.25 (0.83)	17.2 (8.1)	15.19	1.15 2.46	0.92	0.91

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.4. Partial least squares statistics for biological age of 'fresh' Snapper otoliths all ages using a generic NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (months)	LV	R^2	RMSE (months)	Bias (months)	SDR	r	Slope
Sunshine Coast Qld, Winter & Spring 2012		109 (1)	36 – 301 (47.4)	3	0.78	20.6	-0.24	2.31	0.89	0.79
	Sunshine Coast Qld, Winter & Spring 2012	107	37 – 217 (41.5)	3	0.79	19.0	-0.78	2.19	0.89	0.81
Gulf St Vincent SA, Summer & Autumn 2012		147	37 – 267 (51.3)	2	0.91	15.2	-0.05	3.39	0.96	0.91
	Gulf St Vincent SA, Summer & Autumn 2012	151 (1)	37 – 255 (48.0)	2	0.91	14.8	-1.27	3.24	0.95	0.90
Gulf St Vincent SA, Winter & Spring 2012	Cross Validation	93 (1)	66 – 295 (58.1)	2	0.93	15.2	-0.10	3.64	0.96	0.93
Gulf St Vincent SA, All months 2012		186	37 – 295 (58.0)	2	0.94	15.0	-0.08	3.82	0.97	0.93
	Gulf St Vincent SA, All months 2012	216	37 – 267 (47.0)	2	0.89	15.8	3.40	3.08	0.95	0.92
Qld & SA combined, All months 2012		306 (1)	36 – 301 (53.1)	2	0.88	18.4	-0.06	3.87	0.94	0.88
	Qld & SA combined, All months 2012	302	27 – 260 (46.7)	2	0.85	17.9	-0.92	2.61	0.92	0.87

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.5 Partial least squares statistics for biological age of ‘fresh’ Snapper otoliths <156 months using a generic NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (months)	LV	R^2	RMSE (months)	Bias (months)	SDR	r	Slope
Sunshine Coast Qld, Winter & Spring 2012		99 (1)	36 – 145 (32.3)	3	0.79	14.8	-0.14	2.19	0.89	0.80
	Sunshine Coast Qld, Winter & Spring 2012	91	48 – 145 (26.9)	3	0.69	14.9	2.28	1.80	0.83	0.71
Gulf St Vincent SA, Summer & Autumn 2012		130	37 – 148 (27.0)	2	0.85	10.4	-0.02	2.59	0.92	0.85
	Gulf St Vincent SA, Summer & Autumn 2012	115	37 – 145 (29.1)	2	0.85	11.2	-1.18	2.61	0.92	0.84
Gulf St Vincent SA, Winter & Spring 2012	Cross Validation	69	65 – 150 (32.0)	2	0.90	10.3	0.01	3.12	0.95	0.91
Gulf St Vincent SA, All months 2012		148	37 – 150 (31.3)	2	0.89	10.2	-0.02	3.07	0.95	0.90
	Gulf St Vincent SA, All months 2012	164	40 – 146 (28.0)	2	0.84	11.0	-1.40	2.54	0.92	0.83
Qld & SA combined, All months 2012		254	36 – 150 (28.0)	2	0.80	12.6	-0.01	2.22	0.89	0.80
	Qld & SA combined, All months 2012	251	37 – 146 (30.9)	2	0.82	13.0	-2.15	2.37	0.91	0.85

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.6 Partial least squares statistics for biological age of ‘fresh’ Snapper otoliths ≥156 months using a generic NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) in months	LV	R^2	RMSE in months	Bias (months)	SDR	r	Slope
Gulf St Vincent, SA All months 2012	Cross validation	76 (3)	157 – 295 (38.9)	4	0.81	16.7	0.01	0.33	0.90	0.83

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Table A3.7. Partial least squares statistics for biological age of 'fresh' Snapper otoliths of all ages using a geographic location sample specific NIRS wavelength selection

Calibration Set	Validation Set	Sample Size (OR)	Age Range (SD) (months)	LV	R^2	RMSE (months)	Bias (months)	SDR	r	Slope
Sunshine Coast Qld, Winter & Spring 2012		110	36 – 301 (47.4)	2	0.77	20.4	-1.21	2.33	0.88	0.76
	Sunshine Coast Qld, Winter & Spring 2012	107	37 – 217 (41.5)	2	0.79	19.0	-1.67	2.19	0.89	0.80
Gulf St Vincent SA, Summer & Autumn 2012		147	37 – 267 (51.3)	2	0.91	15.4	-0.06	3.33	0.95	0.91
	Gulf St Vincent SA, Summer & Autumn 2012	151	37 – 255 (48.0)	2	0.91	14.8	-0.79	3.26	0.95	0.91
Gulf St Vincent SA, Winter & Spring 2012	Cross Validation	93 (1)	66 – 295 (58.1)	3	0.92	16.1	-0.37	3.55	0.96	0.94

SD = Standard Deviation; OR = Outliers Removed; LV = Latent Variables; RMSE = Root Mean Square Error; SDR = Standard Deviation Ratio

Appendix 4: Guidelines for the Application of Near Infrared (NIR) Spectroscopy for Predicting the Age of Whole Dry Otoliths

This document provides a brief set of guidelines to assist potential users with the basis of developing their own protocols and calibration models for applying NIRS to predict the age of fish based on otoliths. It includes key points about otolith sample presentation and NIRS calibration model development. The chemometric analysis of NIR spectra is complex and it is strongly recommended that NIRS programs to age fish provide relevant training and seek expert guidance in spectral multivariate data analysis techniques and interpretation.

NIR Background

All organic matter is composed of molecules which consist of atoms and groups of atoms which are linked together in various combinations mainly by covalent bonds. All molecules continually vibrate at specific frequencies. Irradiation of molecules by an energy source such as NIR light with wavelength region ranging from 780 - 2500 nm (12820 - 4000 cm^{-1}) causes some molecules to change their vibrations from one energy level to another. When these transitions occur, energy is absorbed at a certain frequency coinciding with those of the molecular grouping in the scanned material. This absorption of energy is detected by NIRS instruments. Molecular vibrations within the spectral region are the result of combinations and overtones of the fundamental vibrations of carbon-hydrogen (C-H), oxygen-hydrogen (O-H) and nitrogen-hydrogen (N-H) that absorb at characteristic wavelengths. NIR spectroscopic measurements obtain information about the relative proportions of these fundamental absorbers which are also repeated throughout the NIR spectral region as overtones or ripples of the fundamental absorber. Therefore, in this case, the chemistry of the otolith provides the specific spectral information that is assumed to be related to fish age.

NIRS requires reference techniques (i.e., visual age estimate of otoliths) to build up calibration routines and to guarantee the proper maintenance of an established calibration with reference to outlier detection and troubleshooting. As a secondary method of determination, the major limitation of NIRS analysis remains its dependence on the accuracy of the reference method. Errors in the visual age assessment of otoliths will perpetuate through NIRS calibration and predictive models. In short, the more accurate and precise the visual age estimate of a species (by geographic location and/or season), the more accurate and precise the NIR calibration and prediction models.

Equipment

The NIR instrument used in the spectral assessment of Barramundi and Snapper otoliths was a Bruker Multi-Purpose Analyser (MPA, Figure 1). The integrating sphere in a reflectance mode configuration in the 12500 - 3600 cm^{-1} (800 - 2780 nm) range was used with spectra averaging (number of scans per sample) of 16 scans per second and resolution of 8 cm^{-1}.

Sample presentation

Dry whole otoliths were placed on the integrating sphere with an upward concave orientation i.e., with the convex orientation exposed to the NIR light (Figure 2). This orientation was found to provide a suitable NIR spectrum. Although, there was no significant difference found in this trial on the rotation of the concave otolith from 0^o, 45^o, 90^o, 180^o on the integrating sphere of the Bruker MPA, it is recommended that a consistent orientation be maintained to reduce possible discrepancies.

Figure 1. MPA instrument.

Figure 2. Concave presentation of an otolith on the integrating sphere window.

Preliminary results suggested no significant spectral differences between otoliths obtained from the left and right side of the fish (i.e., no asymmetry effects), implying that either otolith of an individual can be used interchangeable in the generation of NIRS calibration models. However, we recommend where possible, the consistent use of the left or right otolith for calibration model development and future assessment.

Modern NIRS instruments carry diagnostic software that allows the operator to monitor the performance of the instrument. The diagnostics should be run routinely each day to check whether changes take place during continuous operation. Temperature, relative humidity (RH) of the working environment, and dust all exert influences on the instruments components and performance. Calibration models should ideally be developed over a period of time that exposes the instrument to the most likely fluctuations of these environmental influences.

Various errors can occur in NIRS procedures which decrease the accuracy of prediction. Examples include: technician error in sample preparation, temperature differences in instrument or standards while collecting data, instrument noise and drift, calibration standard instability, changes in instrument wavelength setting, stray light effects, nonlinearity, particle size differences, colour differences with concentration, solvent interaction differences with changing concentration, and the reference method not measuring the same component as the spectroscopic method.

Water is an extremely strong absorber in the NIR spectral region displaying specific bands at 1940 nm (combination), and 1450 nm (first overtones of the O-H stretch), 1190 nm (combination), 970 nm (second overtones of the O-H stretch), and 760 nm (third overtones of the O-H stretch) at 20^{o}C. These bands are subject to shifts as a result of variations in temperature and in hydrogen bonding when water is in organic matrix such as otoliths. As temperature increases, the extent of the hydrogen bonds decreases. Spectrally, the effect of higher temperatures is manifested as a band shift towards lower wavelengths and decrease in bandwidth.

When samples that have different temperatures from that of the calibration sample set are predicted, a bias occurs. It may therefore be better to control the sample temperature before making NIR measurements. Alternatively, temperature variation can be incorporated into the calibration model. In the case of fish otoliths, it is recommended that all otolith samples are allowed to equilibrate to a standard temperature (e.g. room temperature) before NIR spectra are acquired for calibration or prediction.

NIR spectra acquisition

In the current project, labour was required to remove the otoliths from their packaging (i.e., paper envelope or screw-top plastic vial) and place a single otolith onto the integrating sphere of the Bruker MPA. The loading process took about five seconds per otolith. The collection of NIR spectra, with a configuration of 8 cm^{-1} resolution and an average number of scans (16), took about 10 seconds per otolith. Following spectrum collection, the otolith was removed from the integrating sphere and placed back into its original packaging (~five seconds) and the process repeated. Thus, total sample spectrum collection time was about 20 seconds per sample.

The Bruker MPA can be utilised with a carousel (i.e., a sample wheel) that holds up to 30 vials and automatically scans each vial containing the sample for spectra collection. A carousel was not used in the current project as the otoliths moved in the vials when the carousel automatically positioned itself over the integrating sphere. Alternative sample holders could be potentially adapted especially for fish otoliths, allowing 30 otoliths to be loaded onto a carousel and then having the spectrophotometer then automatically scans each otolith unattended, thus saving time and money.

Whatever the method of otolith placement over the NIRS collection window (i.e., integrating sphere), otolith placement optimisation need to be considered to ensure a consistent presentation of the otolith surface to the spectrophotometer with minimal movement at the time of scanning.

Calibration model development

Robust calibration model development and transfer is a major aspect of progressing the adoption of NIR technology by various sectors for the non-invasive assessment of products in a laboratory, in-line and/or in field setting. Building a robust calibration model is a time-consuming and laborious procedure. The dependence on the time-consuming and laborious calibration procedures and the complexity in the choice of data treatment are the main disadvantages of NIR spectroscopy. The development of robust calibration models requires training sets that cover variables such as seasonal otolith growth, yearly (temporal) differences, fish size, diversified age structure and location variables, and measurement conditions (sample handling and presentation). We recommend incorporating geographic variation, species variation, different processing methods and storage condition variation in any further investigation of the stability of NIR spectra from otolith collections.

Calibration maintenance and associated costs is an ongoing process to ensure the robustness of the calibration model so the models continue to accurately predict future unknown samples. Calibration model maintenance includes samples with different biological variability outside the calibration model that may need to be addressed to ensure future predictive performance. In the case of fish otoliths, this may include samples from different months, years, or geographic locations. Calibration development is a regression modelling procedure that identifies the minimum subset of wavelength terms that best explains the chemical property across a population of similar commodities showing multivariate changes in the composition.

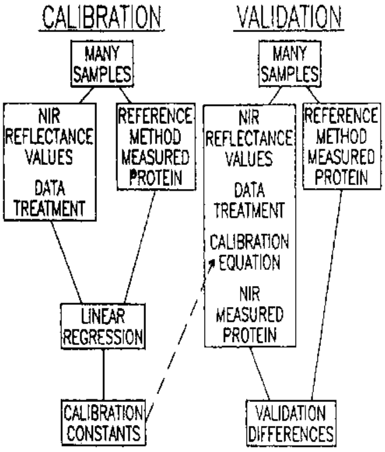


Figure 3. Flow Diagram of calibration and validation steps. From Hruschka (1987).

Calibration Steps:

1. Selection of calibration samples. Ideally, the calibration set must cover an entire constituent range. There are many arguments whether the sample distribution should be uniform or in a Gaussian (normal) fashion with respect to any compositional or functional parameter. If the samples are distributed in Gaussian fashion then the results of the prediction will regress towards the mean. This will persist for the calibrations and cause the NIR-predicted results, and the results at the low end to appear higher, and the high end to appear lower than the reference data. The effect can be minimised by assembling samples to provide an even distribution across the reference range. For NIR application to commercial operation, there is an argument in favour of the Gaussian distribution because the majority of the population of incoming samples will differ from the mean by only one standard deviation value and therefore the Gaussian selection will favour accuracy of most (68%) of the samples (Williams 2013b). A good calibration set should consist of a considerable number of samples. For instance, calibration models in the current study had a target of 10 samples per age class for each geographic location and season. However, it is more likely in a fisheries sampling program to have a target of 10 otoliths per length class, particularly if the aim of estimating fish age is to develop an age-length key that is applied to a greater number of fish who were sampled for length but not otoliths/age.

2. Selection of a suitable reproducible sample preparation and presentation method (i.e., dried whole otolith samples are placed on the integrating sphere with an upward concave orientation with the convex orientation being exposed to the NIR light). Post-mortem handling practices should be carefully evaluated for potential effects on the chemical composition of otoliths and where possible standardised to ensure any effects are consistent across calibration and prediction sample sets.

3. Recording of the NIR spectra by subjecting the calibration samples to analysis by the NIR instrument at the required wavelengths.

4. Accurate analysis of the sample against the appropriate reference method for the constituent of interest. The quality of the reference method value dictates that of the calibration model.

5. Development of a calibration equation using statistical modelling techniques such as multiple linear regression (MLR) and partial least squares (PLS) to establish a correlation between the NIR spectral data and the chemical parameters of the sample set. The calibration coefficients define the weights given to the different wavelengths in the linear equation, which is a simple formula that will therefore be used as a common way of expressing the analytical results from all the different linear calibration methods treated.

6. Validation of the model to ensure that the model accurately predicts the property of interest in samples not subjected to the calibration process.

7. Predicting unknown samples. A robust, accurate model can be used to predict rapidly the property of interest in new, unknown samples.

For a more in-depth guide on calibration development and evaluation methods refer to Williams (2013):

   a) Calibration development and evaluation methods - A. Basics http://www.impublications.com/subs/nirn/v24/N24_0524.pdf;
   b) Calibration development and evaluation methods – B. Set-up and evaluation http://www.impublications.com/subs/nirn/v24/N24_0620.pdf.

Data analysis (Chemometrics)

NIR spectra contain a great deal of physical and chemical information about molecules of the item scanned. However, this information cannot always be extracted straightforwardly from the spectra as firstly, the NIR spectrum consist of a number of bands arising from overtones in combination modes that overlap heavily with each other. Secondly, NIRS deals quite often with "real-world" samples, which yield rather poor signal-to-noise (S/N) ratio, baseline fluctuations, and severe overlapping of bands due to various components. Chemometrics have been employed to extract rich information from NIR spectra and a main part of chemometrics is multivariate data analysis.

In the current project data analysis used "The Unscrambler" Software Version 9.8 (Camo, Oslo, Norway) for multivariate analysis of NIR spectral data. However, there are numerous commercial chemometric software packages available that are equally suitable.

Model evaluation

Calibration models can be evaluated using cross validation (leave-one-out or segmented) or an independent test set (validation set) of samples. The cross validation method is recommended for small samples ($n<100$), while an independent test set is recommend for larger sample sets ($n>100$). In the cross validation process, a single sample or group of samples for the segmented case is withdrawn from the total set with the calibration being developed on the remaining samples. With regards to independent validation sets, both the calibration and validation set should contain a similar distribution of reference data.

For the current project, partial least squares (PLS) regression was used to build the otolith-age prediction models based on the diffuse reflectance spectral data. Before the development of a calibration model, the variation of the spectral data needs to be investigated by principal component analysis (PCA) and obvious atypical spectra eliminated. PLS regression attempts to establish a correlation between the spectral data and the otolith age reference data set (based on standard visual assessment of sectioned otoliths) in order to find the optimal model. In other words, the calibration equation of the NIR and chemical loadings combine mathematically to yield the calibration model which is then used for analysis of future unknown samples. Prior to analysis, raw spectral data was mathematically transformed to remove defects observed in the spectra (noise, base line drift etc.).

Wavelength selection

Areas in the spectrum where maximum differences occur in response to changes in concentration of substances of interest can indicate that these spectral regions or wavelengths will likely be used in an accurate analysis. Spectral region selection can significantly improve the performance over full-spectrum calibration techniques and maybe one of the most critical aspects of NIR analysis. Specific spectral regions are selected where co-linearity is not so important, generating more stable models with superior interpretability. In practice, this requires the identification of a subset of the full spectrum that will produce the lowest prediction error.

Model robustness and variance

Robust calibration model development is a major aspect of NIR technology. Unfortunately, while NIRS has many merits, building calibration models is a time-consuming procedure, and the calibration model may provide unsatisfactory prediction results, because either or both NIR instruments and samples are contaminated over time.

Predictive calibration models may change or lose robustness because of instrument replacement; instrument drift or shift (change in response by the instrument due to wear, replacement of vital parts etc.); differences between NIR instruments; change in the sampling or measuring environment (e.g., temperature, humidity); or a change in the sample physical or chemical constitution (particle size, surface texture).

The major challenge is to ensure that the calibration is robust, that is, that the calibration holds across seasonal otolith growth, temporal differences and geographic location variables. These variables can impact on the otolith chemistry and effect NIR predictions. Sources of variance therefore need to be determined and included into the calibration model. Most often the sample composition is considered, but any factors that affect spectral characteristics including operating conditions need to be considered. Thus any factors that represent changes in the processing conditions should be included.

Temporal and spatial effects have major impacts on the robustness of otolith calibration models. However, this variability can be accommodated in the calibration models by combining samples from different locations and including samples with a wide range of temporal and seasonal differences (i.e., incorporating numerous seasons and years). Thus, including a wide range of biological variability in the calibration model enables robust prediction of future samples. This result is not an unusual result in NIRS work on horticultural commodities, where biological properties of produce often vary spatially and seasonally. In otoliths, differences in geographic specific NIRS calibration models probably reflect spatial differences in the chemistry of otoliths. Geographic variability may be able to be addressed in some cases through applying a bias correction, when bias is consistent across the sample set.

The development of robust calibration models requires training sets that cover variables such as seasonal otolith growth, yearly (temporal) differences, fish size, diversified age structure and location variables, and measurement conditions (sample handling and presentation). However, in some cases, incorporation of more biological variability (at the risk of including atypical data) in the calibration set can significantly reduce the models prediction accuracy.

In general, calibration model error may easily double when a calibration model is applied to a spectral data set of a different season, year (temporal) or geographic location. This lack of robustness often translates into bias. Prediction bias for new populations can be corrected by model updating or direct bias adjustment.

Model performance

Several statistical criteria are used to describe the performance of NIR spectroscopy models to predict the required chemical parameter of the unknown sample, these include:

- The ***coefficient of determination*** (R^2), which in the Unscrambler chemometric software package is the proportion of the total variance accounted for by the explained variance for the given number of latent variables. This parameter has values between 0 and 1e.
- The ***root mean standard error of calibration*** (RMSEC) is an estimation of the variation of the reference and predicted values of the calibration population, that is, how well the calibration model fits the calibration set.
- The ***root mean square error of prediction*** (RMSEP) is used to indicate the predictive performance of the calibration model, that is, an estimation of the variation of the reference and predicted values of the validation population.
- For ***root mean square error of cross validation*** (RMSECV), a leave-one-sample-out cross validation is performed: the spectrum of one sample of the training set is deleted from this set and a PLS model is built with the remaining spectra of the training set. The left out sample is predicted with this model and the procedure is repeated with leaving out each of the samples of the training set. Cross validation of a calibration model makes it possible to select the optimum number of latent variables or factors, that is, the number giving the minimum prediction error for the calibration set.

- The '***bias***' corresponds to the average difference between measured and predicted values. If there is no such difference, the bias will be zero. If the bias values are negligible, Buning-Pfaue (2003) suggest that the standard error of prediction (SEP) value can be equated with the standard deviation (SD) and therefore, in the case of a statistical certainty of 0.95, the maximum error range can be specified as ±2 (1.96) SEP, called the 'maximum error range'. A ***bias adjustment*** is appropriate when all predicted values consistently read higher (or lower) than the measured (reference) values. A large positive (negative) bias can be corrected by adding (subtracting) the **absolute** bias value to all reference values and re-running the prediction model.
- The ***standard deviation ratio*** (SDR) is the ratio of the standard deviation (SD) of the population divided by the RMSEP or the RMSECV. The SDR statistic enables comparison of model performance across populations with different SD. In theory (for large normally distributed data sets) SDR is equal to, and indicates more directly than either R^2 or RMSEP separately can, the relative predictive performance of a model; the higher the value, the greater the power. McGlone and Kawano (1998), suggest that an SDR of >3 is adequate to support sorting /grading into 3 classes. While Guthrie *et al.* (1998), suggest that for NIR spectroscopy to be commercially useful in fruit grading, the technique must be capable of sorting fruit into at least two grades (i.e., above and below an acceptable level) with approximately 80% accuracy. This requirement involves attainment of a validation correlation coefficient of at least 0.65. Golic and Walsh (2006), report that an SDR of 2.5 allows sorting into two grades.
- The term ***ratio of*** (standard error of) ***performance deviation*** (RPD) is similar to the SDR except the RPD uses a bias corrected RMSEP or RMSEC. If the SEP is close to the SD of the reference data of the validation samples set (whether cross validation or a test set is used), then the calibration model is not efficiently predicting and is of no practical value in analysis (Baillères *et al.* 2002; Williams 2007). If SEP = SD (i.e. RPD is of 1.0), the calibration is essentially predicting the population mean (Baillères *et al.* 2002; Williams 2007). The interpretation of RPD values for 'difficult' applications, such as forage analysis, high moisture materials, such as fruit, vegetables and meat, soils and manures, where the application of NIR spectroscopy is affected by the more complex nature of the materials (Williams 2007). An RPD below 2 cannot give a relevant prediction, while an RPD value of 2.0 – 2.4 is regarded as adequate for rough screening. RPD values of 2.5 to 2.9 are regarded as fair for screening, 3.0 -3.4 are regarded as satisfactory for quality control, 3.5 to 4.0 very good for process control and 4.1+ are excellent for any application (Williams 2008). The RPD in relation for low moisture products such as grains, flours and meals, for composition in comparison to high moisture products has a different interpretation in terms that the statistical consequence of the lower values is not changed, but it places them in a category of explanation that is more realistic in terms of likely to be achieved with these applications (Williams 2008). For example, an RPD below 2 cannot give a relevant prediction, while an RPD value of 2.0 - 3.0 is regarded as adequate for rough screening. RPD values of 3.0 and above are regarded as satisfactory for screening (Williams 2007), values of 5 and above are suitable for quality control analysis, and values of above 8 are excellent, and can be used in any analytical situation (Baillères *et al.* 2002).

All calibrations should be examined by analysis of unknown samples. The lowest SEP and highest R^2 will be the optimum combination of wavelength and pre-processing treatment. Ideally, the SEP and the bias should both be equal to zero with a small difference between SEC (standard error of calibration) and SEP (Williams 1987; Gomez *et al.* 2006). Thus, a high R^2 with a low SEP and bias means that the NIR results are accurate over the anticipated range, and likely to remain so, provided the statistics were based on a sufficient number of observations.

The R_c^2 (calibration model coefficient of determination) is a function of SD and RMSEC, and the R_v^2 (validation coefficient of determination) is a function of SD and RMSEP (when bias = 0) (Eqn. 1). The R^2 can be improved by increasing the SD of the calibration population. Therefore, an evaluation of a model using the R^2 statistic should be considered in conjunction with knowledge of the SD (which should be equivalent to that of the population to be predicted).

$$R^2 = 1 - (SEP/SD)^2 \quad \text{Eqn. 1}$$

A low R^2 between NIR and reference data means that the NIR analysis has not been successful, and it is possible that NIRS may not be applicable to that particular analysis. In general, the SEC decreases as the correlation coefficient (r) increases; and provided the r and R^2 are high (above 0.9), the accuracy can always be fine-tuned by a slope/bias adjustment, despite an apparently high initial SEP (Hruschka 1987; Williams 1987). If the r is 0.8 or below, the SEP maybe capable of improvement Williams (1987), however, it is not possible to obtain really high accuracy unless the r is high (≥0.90).

Latent variables and over fitting

As more latent variables (or factors) are included in the calibration model, the model begins to fit the random errors embedded in the spectra and concentrations. Therefore, the RMSEC will always decrease as more factors are added and may end up much lower than the laboratory error of the reference method (Boyworth and Booksh 2001). When extra factors that mostly describe random errors are included in the calibration model, these factors will not fit the errors in the future samples and the RMSECV and RMSEP may increase. The best policy is to keep the number of wavelength terms as low as practicable. A simple rule of thumb proposed by Hruschka (1987), is to use 5 to 15 samples for each regression and data treatment constant and for any parameter of the data treatment (such as wavelength) that is allowed to vary. Lammertyn *et al.* (2000), applies a similar rule of thumb, where the ratio of the number of samples to the number of variables should be equal to or larger than 10. While Williams (2013a) recommends to assemble 20 to 25 samples for each factor that you intend to use in development of a calibration; for example: if 15 PLS factors are specified, at least 20 samples should be assembled per factor, for MLR calibrations, 25 samples per wavelength is recommended.

Appendix 5: References

Baillères H, Davieux F, Ham-Pichavant F (2002) Near infrared analysis as a tool for rapid screening of some major wood characteristics in eucalyptus breeding program. *Annals of Forrest Science* **59**, 479-490.

Beamish RJ, Fournier DA (1981) A method for comparing the precision of a set of age determinations. *Canadian Journal of Fisheries and Aquatic Sciences* **38**, 982-983.

Borelli G, Guibbohni ME, Mayer-Gostan N, Priouzeau F, De Pontual H, Allemand D, Puverel S, Tambutte E, Payan P (2003) Daily variations of endolymph composition: relationship with the otolith calcification process in trout. *Journal of Experimental Biology* **206**, 2685-2692.

Boyworth MK, Booksh KS (2001) 'Aspects of multivariate calibration applied to near-infrared spectroscopy. .' (Marcel Dekker Inc, New York, USA and Basel, Switzerland.)

Buning-Pfaue H (2003) Analysis of water in food by near infrared spectroscopy. *Food Chemistry* **82**, 107–115.

Campana SE (1999) Chemistry and composition of fish otoliths: pathways, mechanisms and applications. *Marine Ecology Progress Series* **188**, 263-297.

Campana SE (2005) Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of Fish Biology* **59**, 197-242.

Campana SE, Annand MC, McMillan JI (1995) Graphical and statistical methods for determining the consistency of age-determinations. . *Transactions of the American Fisheries Society* **124**, 131-138.

Chang M-Y, Geffen AJ (2013) Taxonomic and geographic influences on fish otolith microchemistry. *Fish and Fisheries* **14**, 458-492.

Chilton DE, Beamish RJ (1982) Age determination methods for fishes studies by the Groundfish Program at the Pacific Biological Station. *Canadian Journal of Fisheries and Aquatic Sciences* **60**, 1-102.

Craine M, Rome B, Stephensom P, Wise B, Gaughan D, Lenanton R, Steckis R (2009) 'Determination of a cost effective methodology for ongoing age monitoring needed for the management of scalefish fisheries in Western Australia. Final FRDC Report - Project 2004/042.' Department of Fisheries, Western Australia.

Fairclough DV, Molony BW, Crisafulli BM, Keay IS, Hesp SA, Marriott RJ (2014) 'Status of demersal finfish stocks on the west coast of Australia.' Department of Fisheries, Western Australia.

Ferrell DJ, Henry GW, Bell JD, Quartararo N (1992) Validation of annual marks in the otoliths of young snapper, *Pagrus auratus* (Sparidae). *Australian Journal Marine and Freshwater Research* **42**, 1051-1055.

Fisheries Queensland (2010) 'Fisheries Long Term Monitoring Program Sampling Protocol – Barramundi (2008 onwards) Section 1.' Department of Employment, Economic Development and Innovation, Brisbane, Australia.

Forrest AJ, Pooley SG, Exley P, Mayze J, Paulo C (2012) 'Management of 'tough fish syndrome' in tropical Saddletail snapper (*Lutjanus malabaricus*) to re-instill market confidence.' Department of Agriculture, Fisheries and Forestry, Brisbane.

Fowler AJ, McGarvey R, Burch P, Feenstra JE, Jackson WB (2010) 'Snapper (*Chrysophrys auratus*) Fishery. Fishery Assessment Report to PISA.' South Australian Research and Development Institute (Aquatic Sciences), Adelaide.

Francis RICC, Campana SE (2004) Inferring age from otolith measurement: a review and a new approach. *Canadian Journal of Fisheries and Aquatic Sciences* **61**, 1269-1284.

Francis RICC, Harley SJ, Campana SE, Doering-Arjes P (2005) Use of otolith weight in length-mediated estimation of proportins at age. *Marine and Freshwater Research* **56**, 735-743.

Francis RICC, Paul LJ, Mulligan KP (1992) Ageing of adult snapper (*Pagrus auratus*) from otolith annual ring counts: validation by tagging and oxytetracycline injection. *Australian Journal Marine and Freshwater Research* **43**, 1069-1089.

Gauldie RW, Thacker CE, West IF, Wang L (1998) Movement of water in fish otoliths. *Comparative Biochemistry and Physiology Part A* **120**, 551-556.

Genstat (2013) Genstat for Windows, Release 16.1.0.10916, . In. (VSN Internatinal Ltd: Oxford, UK)

Gillanders BM (2002) Temporal and spatial variability in elemental composition of otoliths: implications for determining stock identity and connectivity of populations. *Canadian Journal of Fisheries & Aquatic Sciences* **59**, 669-679.

Golic M, Walsh KB (2006) Robustness of calibration models based on near infrared spectroscopy for the in-line grading of stonefruit for total soluble solids content. *Analytica Chemica Acta* **555**, 286-291.

Gomez AH, He Y, Pereira AG (2006) Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *Journal of Food Engineering* **77**, 313-319.

Guthrie J, Wedding B, Walsh K (1998) Robustness of NIR calibrations for soluble solids in intact melon and pineapple. *Journal of Near Infrared Spectroscopy* **6**, 259-265.

Hale R, Swearer SE (2008) Otolith microstructural and microchemical changes associated with settlement in the diadromous fish Galaxias maculatus. *Marine Ecology Progress Series* **354**, 229-234.

Halliday I, Robins J (2007) 'Environmental flows for sub-tropical estuaries: understanding the freshwater needs of estuaries for sustainable fisheries production and assessing the impacts of water regulation. .' Queensland Department of Primary Industries,, Brisbane.

Halliday IA, Saunders T, Sellin MJ, Allsop QA, Robins JB, McLennan M, Kurnoth P (2014) 'Flow impacts on estuarine finfish fisheries of the Gulf of Carpentaria. .' Queensland Department of Agriculture, Fisheries and Forestry, Brisbane.

Hedges KJ, Ludsin SA, Fryer BJ (2004) Effects of ethanol preservation on otolith microchemistry. *Journal of Fish Biology* **64**, 923-937.

Hruschka WR (1987) 'Data analysis: wavelength selection methods.' (The American Association of Cereal Chemist, Inc, Minnesota USA.)

Hunt R (1977) Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics* **42**, 501-513.

Kawano S (2002) ' Sampling and sample presentation.' (Wiley-VCH Verlag GmbH, Weinheim, Germany.)

Lammertyn J, Peirs A, De Baerdemaeker J, Nicolai B (2000) Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. *Postharvest Biology and Technology* **18**, 121-132.

Lepak JM, Cathcart N, Hooten MB (2012) Otolith mass as a predictor of age in kokanee salmon (*Ocorhynchus nerka*) from four Colorado reservoirs. *Canadian Journal of Fisheries and Aquatic Sciences* **69**, 1569-1575.

Lou DC, Mapstone BD, Russ GR, Begg GA, Davies CR (2007) Using otolith weight-age relationships to predict age based metrics of coral reef fish populations across different temporal scales. *2007* **83**, 216-227.

Marine-Institute (2007) 'OTO - Otolith Training Online.'

McDougall A (2004) Assessing the use of sectioned otoliths and other methods to determine the age of centropomid fish, barramundi (*Lates calcarifer)* (Bloch), using known-age fish. *Fisheries Research* **67**, 129-141.

McGlone VA, Kawano S (1998) Firmness, dry-matter, and soluble-solids assessment of postharvest kiwi fruit by NIR spectroscopy. *Postharvest Biology and Technology* **13**, 131-141.

Milton DA, Chenery SR (1998) The effect of otolith storage methods on the concentrations of elements detected by laser-ablation ICPMS. *Journal of Fish Biology* **53**, 785-794.

Moron A, Cozzolino D (2003) Exploring the use of near infrared reflectance spectroscopy to study physical properties and microelements in soils. *Journal of Near Infrared Spectroscopy* **11**, 145-154.

Ochwada FA, Scandol JP, Gray CA (2008) Predicting the age of fish using general and generalized linear models of biometric data: A case study of two estuarine finfish from New South Wales, Australia. *Fisheries Research* **90**, 187-197.

Owen JF (1998) Re-interpretation of remains of snapper (Pagrus auratus) from Holocene middens at Bass point and Currarong,. *Proceedings of the Linnean Society of New South Wales* **120**, 191.

Parmentier E, Cloots R, Warin R, Henrist C (2007) Otolith crystals (in Carapidae): Growth and habit. *Journal of Structural Biology* **159**, 462-473.

Payan P, Edeyer A, De Pontual H, Borelli G, Boeuf G, Mayer-Gostan N (1999) Chemical composition of saccular endolymph and otolith in fish inner ear: lack of spatial uniformity. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology* **277**, R123-R131.

Proctor CH, Thresher RE (1998) Effects of specimen handling and otolith preparation on concentration of elements in fish otoliths. *Marine Biology* **131**, 681-694.

Radtke RL, Shafer DJ (1992) Environmental sensitivity of fish otolith microchemistry. *Australian Journal of Marine and Freshwater Research* **43**, 935-951.

Rigby CL, Wedding BB, Grauf S, Simpfendorfer CA (2014) The utility of near infrared spectroscopy for age estimation of deepwater sharks. *Deep Sea Research Part I: Oceanographic Research Papers* **94**, 184-194.

Rigby CL, Wedding BB, Grauf S, Simpfendorfer CA (In review (a)) A novel method for aging sharks using Near Infrared Spectroscopy. *Marine and Freshwater Research*.

Robertson SG, Morison AK (1999) A trial of artificial neural networks for automatically estimating the age of fish. *Marine and Freshwater Research* **50**, 73-82.

Rooker JR, Zdanowicz VS, Secor DH (2001) Chemistry of tuna otoliths: assessment of base composition and post mortem handling effects. *Marine Biology* **139**, 35-43.

Rowell K, Dettman DL, Dietz R (2010) Nitrogen isotopes in otoliths reconstruct ancient trophic position. *Environmental Biology of Fishes* **89**, 415-425.

Shenk JS, Workman JJ, Westerhaus MO (2001) 'Application of NIR spectroscopy to agricultural products.'

Sohn D, Kang S, Kim S (2005) Stock identification of chum salmon (Oncorhynchus keta) using trace elements in otoliths. *Journal of Oceanography* **61**, 305-312.

Staunton-Smith J, Robins JB, Mayer DG, Sellin MJ, Halliday IA (2004) Does the quantity and timing of fresh water flowing into a dry tropical estuary affect year-class strength of barramundi (*Lates calcarifer*)? *Marine and Freshwater Research* **55**, 787-797.

Stuart IG, McKillup SC (2002) The use of sectioned otoliths to age barramundi (*Lates calcarifer*) (Bloch, 1790) [Centropomidae]. *Hydrobiologia* **479**, 231-236.

Swan SC, Gordon JDM, Shimmield T (2003) Preliminary Investigations on the Uses of Otolith Microchemistry for Stock Discrimination of the Deep-water Black Scabbardfish (Aphanopus carbo) in the North East Atlantic *Journal of Northwest Atlantic Fishery Science* **31**, 221-231.

Tabouret H, Lord C, Bareille G, Pecheyran C, Monti D, Keith P (2011) Otolith microchemistry in Sicydium punctatum: indices of environmental condition changes after recruitment. *Aquatic Living Resources* **24**, 369-378.

Thomas D, McGoverin C, Chinsamy A, Manley M (2011) Near infrared analysis of fossil bone from the Western Cape of South Africa. *Journal of Near Infrared Spectroscopy* **19**, 151-159.

Thresher RE (1999) Elemental composition of otoliths as a stock delineator in fishes. *Fisheries Research* **43**, 165-204.

Wedding BB, Forrest AJ, Wright C, Grauf S, Exley P (2014) A novel method for the age estimation of Saddletail snapper (*Lutjanus malabaricus*) using Fourier Transform-near infrared (FT-NIR) spectroscopy. *Marine and Freshwater Research* **65**, 894-900.

Williams P (2007) Grains and seeds. . In 'Near-infrared spectroscopy in food science and technology '. (Eds Y Ozaki, WF McClure and AA Christy) pp. 165-218. (John Wiley & Sons, Inc., New Jersey, United States of America.)

Williams P (2008) 'Near-Infrared Technology - Getting the best out of light ' (PDK projects, Inc.: Nanaimo, Canada)

Williams P (2013a) Tutorial: Calibration development and evaluation methods A. Basics. *NIR news* **24**, 24.

Williams P (2013b) Tutorial: Calibration development and evaluation methods B.Set-up and evaluation. *NIR news* **24**, 20-22.

Williams PC (1987) Variable affecting near-infrared reflectance spectroscopic analysis. In 'Near-infrared technology in the agricultural and food industries'. (Eds P Williams and K Norris) pp. 143-167. (The American Association of Cereal Chemist, Inc.: St Paul, Minnesota USA.)

Worthington DG, Fowler AJ, Doherty PJ (1995) Determining the most efficient method of age determination for estimating the age structure of a fish population. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 2320-2326.

Zorica B, Sinovcic G, Kec VC (2010) Preliminary data on the study of otolith morphology of five pelagic fish species from the Adriatic Sea (Croatia). *Acta Adriatica* **51**, 89-96.