Spring 2021

Exploring Effects of Sample Storage, Preparation, and Tissue Type On Fourier Transform-Near Infrared Spectroscopy (FT-NIRS) Ageing Across Fish Taxa

Keith Fuller

EXPLORING EFFECTS OF SAMPLE STORAGE, PREPARATION, AND TISSUE TYPE ON FOURIER TRANSFORM-NEAR INFRARED SPECTROSCOPY (FT-NIRS) AGEING ACROSS FISH TAXA

by

Keith Fuller

Bachelor of Science
Wright State University, 2015

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Biological Sciences

College of Arts and Sciences

University of South Carolina

2021

Accepted by:

Joseph Quattro, Major Professor

Joshua Stone, Committee Member

Thomas Hilbish, Committee Member

Roger Sawyer, Committee Member

William Driggers III, Committee Member

Tracey L. Weldon, Interim Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

I would like to thank my advisor, Joe Quattro, and members of my committee, Josh Stone, Jerry Hilbish, Roger Sawyer, and Trey Driggers for their continuing support and encouragement throughout this process. I could not have done any of this without their help, and am eternally grateful for having it. My research changed direction multiple times over my time at USC, and Joe was always ready to offer new suggestions and ideas whenever asked. Tori Long and Michelle Passerotti from the Quattro lab were instrumental in this effort as well and have my eternal gratitude.

I would also like to thank everyone who assisted in my search for sample sets, including everyone at SC DNR, Bayless Fish Hatchery, and the Cohen-Campbell Fisheries Center. Particular thanks are due to Lane Hite, Jarrett Gibbons, and Leo Rose for their assistance in obtaining the striped bass larvae and for going out of their way to help with any and all of the numerous questions I had for them. Acknowledgement is also due to Bryan Frazier and all others at the SC DNR Charleston facility, both for lending me the finetooth vertebrae sample set and allowing me out on multiple longline sampling trips. James Sulikowski, at Arizona State University, likewise has my thanks for his lending of a substantial skate sample set and his willingness to help.

Finally, I would like to thank all of my friends and family for their constant support. It is cliché but true to say that I could not have done this without them. My lovely wife, Lydia, deserves special recognition in this multi-year research endeavor, if nothing else for not begrudging my occasional month-long longline trips.

ABSTRACT

The accurate determination of age for fish is a vital part of both population management and ichthyological research. However, the methods which are primarily employed to age fish can be difficult, time consuming, inaccurate, or some combination thereof. Most ageing is currently done via hard parts (such as otoliths, vertebrae, or scales) which are read in a manner similar to tree rings with markings corresponding to known or presumed periods of time. Recently, Fourier transform-near infrared spectroscopy (FT-NIRS) has been investigated as a novel tool to age fish more quickly and objectively. This method works by recording the vibrational frequencies of molecular bonds in a scanned sample which are then correlated with age through partial least squares (PLS) regression models. Several fundamental questions remain before wider usage, however, including questions of sample storage, structure utility, errors introduced by reliance upon traditionally determined ages, and taxa- specific issues of age resolution. In this study, a high degree of ageing accuracy was found with FT-NIRS for *Morone saxatilis* and *Carcharhinus isodon* samples in each storage (frozen or EtOH) and preparation (raw or bleached) method investigated. Ageing was unsuccessful with mounted skate (*Leucoraja ocellata*) vertebral sections. Most samples of *M. saxatilis* were able to be aged to within 2 days of true age, and ~90% were aged to within 5. Over half of the *C. isodon* samples were aged to within 1 year of traditionally determined age. The surprising degree of success found using whole *M. saxatilis* larvae suggests further streamlining potential as well as the possibility of using FT-NIRS for nonlethal ageing in

situ. While unsuccessful at ageing the skate samples tested, the high accuracy, speed, and cross-structure applicability demonstrated here by FT-NIRS strongly justifies continued exploration into its utility within the field of fisheries science.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%RMSE ..... Root Mean Square Error as a Percent of Maximum Value

$\%RMSE_{Range}$ ..... Root Mean Square Error as a Percent of the Range of Data Used

APE .....Average Percent Error

CV ..... Coefficient of Variation

R^2..... Coefficient of Determination

RMSECV .....Root Mean Square Error of Cross Validation

RMSEP ..... Root Mean Square Error of Prediction

RPD..... Residual Prediction Deviation

CHAPTER 1

REVIEW OF FISH AGEING

Effective management of fish populations requires the ability to accurately assess age-at-catch information across ichthyofauna taxa (Campana 2001; Maunder and Punt 2013; Ono et al. 2015). While a wide variety of techniques are used for this, varying greatly between taxa, these techniques are often labor intensive, require specialized training, and are prone to multiple sources of error (Campana 1999; Helser et al. 2019). For commercially valuable populations, accurately determining age-at-catch is a vital part of fishery stock assessments, management plan creation, and predicting the impact of future fisheries activities (Maunder and Punt 2013; Ono et al. 2015). Age-at-catch information directly guides species management plans by indicating the strength of a given year's recruitment and the general population structure, and therefore can impact public policy (Ono et al., 2015). Age at catch data is also directly responsible for such vital information as age at maturity estimates, age at length estimates, and estimates of mortality (Campana and Thorrold 2001; Tahvonen et al. 2018). However, the difficulty and time associated with many traditional ageing methods makes this vital part of population evaluation costly and labor intensive. While different ageing methods vary in expense, all come with tradeoffs in accuracy, time expenditure, and labor required (Campana 2001; Begg et al. 2005). Given the importance of age information across fisheries fields, the continued exploration of ageing methods is unsurprising.

The most commonly used method is hard structure-based ageing. Hard structures, such as otoliths, opercula, scales, vertebrae, or spines, are aged much like a tree, with "rings" corresponding to known periods of time. The difference in structure utilized depends primarily upon taxa and age. While the principle behind this method goes back many years, with references in published literature made as early as the 17th century to the rings in calcified structures of fish and in the late 1800's on the demonstrated accuracy of scale ageing in young carp (*Cyprinus carpio*), advancements have continued to the present day in both method and statistical analysis (Carlander 1987; Kerns and Lombardi-Carlson 2017).

Hard-Structure Ageing

Scales

Scales were the earliest recorded structure to be used for ageing purposes, and are still utilized in numerous taxa (Kerns and Lombardi-Carlson 2017; e.g. Branigan et al. 2019). While the earliest published record utilizing fish scales for age determination comes from the 1800's, some speculate that the idea dates back as far as Aristotle (Jackson 2007). The essential preparation of scales (either for direct examination or impression creation) has varied relatively little in that time. Initially, scales were mounted on slides in a gelatin media before being examined under a microscope or enlarged and projected (Van Oosten 1929). The use of scale impressions on celluloid became more popular in the 1930's, before the development of more convenient plastics in the 1940's (Nesbit 1934; Carlander 1987). Many modern studies still simply examine a scale placed

between 2 glass slides, but impressions in plastic do have a number of benefits (such as durability, ease of storage, and ease of cleaning) (McInerny 2017; e.g. Long et al. 2018).

Ctenoid and cycloid scales contain alternating bands of organic and mineral matrix, with the resultant rings known as circuli (Schonborner et al. 1979; McInerny 2017). These bands reflect periods of growth, with lowered growth resulting in more clustered bands. Once prepared for ageing, scales are examined for patterns known as annuli. Differentiation of annuli can be quite tricky and requires extensive experience and reference to known-age samples, but they can very roughly be defined as patterns of tightly and atypically grouped circuli (Quist et al. 2012; Wright et al. 2015; McInerny 2017). Ideally, each scale forms one annulus per year, which allows the number of annuli present to represent the age of the specimen. While ganoid scales do record age in a manner similar to bones (Buckmesier et al. 2012) with successive layers indicating growth, the standard method of scale ageing is used primarily on ctenoid and cycloid scales.

Numerous problems exist with scale based ageing methods. First, scales can be reabsorbed during periods of stress and calcium deprivation, and can be damaged and removed throughout the course of a fish's life (Whitledge 2017). While the scales that grow back do continue to record periodic growth after they reach the original size of the lost scale, they will no longer accurately reflect the age of the fish (Bereiter-Hahn and Zylberberg 1993). Second, age determination by scales consistently underages older fish across multiple taxa (Beamish and McFarlane 1987). The crowding which occurs after roughly 8 years of age (varying by taxa) largely precludes ageing individuals older than that, and more importantly asymptotic growth in older individuals means that periods of

non-growth will at times exceed 12 months, resulting in increasingly few annuli created as a fish gets older (Beamish and McFarlane 1987). This has historically resulted in significant underageing of numerous taxa (e.g. Power 1978; Beamish and McFarlane 1983). The existence of "false annuli" which are sometimes difficult to distinguish from true annuli and can result from stress or environmental change complicates this further (Meunier 2002; Arechavala-Lopez et al. 2012). Issues of reader interpretation also exist and will be discussed below in the "Errors in Ageing" section. Scales are often considered the structure which requires the most skill and experience to accurately age (McInerny 2017). Overall, in numerous taxa they have been shown to be less accurate than other commonly used ageing structures (McInerny 2017).

Despite these issues, the use of scales for ageing purposes continues. They are the preferred method of ageing species of concern, in which the lethal removal of a number of individuals sufficient to provide population age estimates via otoliths would have an undesirable effect on the population (McInerny 2017). Scales are also used in shorter lived freshwater taxa, where the accuracy is often comparable to otolith ageing but without the lethality of collection (e.g. Beamish and Harvey 1969; Schmitt and Hubert 1982; Niewinski and Ferreri 1999; Schrank and Guy 2002). In some cases, scales have proven to be more accurate than alternative structures for ageing very young fish (Taylor and Weyl 2012).

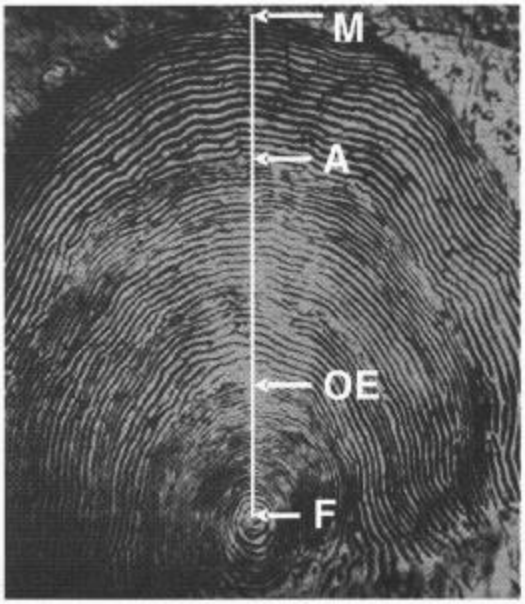


Figure 1.1. Scale of coho salmon (*Oncorhynchus kisutch*), with the annulus marked (A). Each alternating band of light and dark comprises a circuli. The marker (F) represents the initial point of formation, (OE) represents ocean entry, and (M) represents the margin (from Fisher and Pearcy, 2005).

Otoliths

Otoliths are the most common structure used for fish ageing, and arguably the most widely applicable (Secor et al. 1995; Campana 2001; Long and Grabowski 2017). They are only appropriate for use in bony fishes (mostly teleosts), but within this group are widely utilized (Long and Grabowski 2017). Otoliths are found in the inner ear of fish, in one of three possible canals- otoliths from the utriculus are known as lapilli, otoliths from the saccule are known as sagittae, and otoliths from the lagena are known as asterisci. For most teleost taxa, the sagittae are the largest of the otolith varieties, and are the most commonly used for ageing purposes (Long and Grabowski 2017; Kerns and Lombardi-Carlson 2017). However, sagittae are not used for all species, and the need for increased care about otolith description and identification has been discussed (Secor et al. 1992; Long and Stewart 2010; Long and Grabowski 2017).

Otoliths consist mostly of calcium carbonate ($CaCO_3$) with a protein matrix and trace minerals (Campana and Thorrold 2001; Popper et al. 2005). The alternating calcium-rich and protein-rich layers create visually distinctive bands on the otolith, which correspond to periods of growth, seasonality, or diel pattern (Whitledge 2017). Unlike scales, this allows otoliths in some species to be used for both annual and daily age determination (figures 1.2 and 1.3) (Pannella 1971; Long and Grabowski 2017). Annual growth rings are significantly wider than daily growth rings, and can therefore be read without as much magnification (Wright et al. 2002). Otoliths do, however, often require significant preparation to accurately read, and still present issues of interpretation.

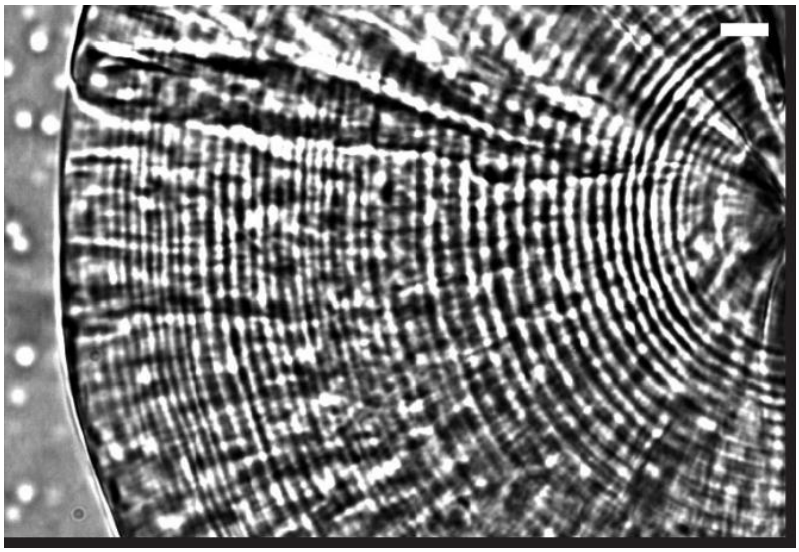


Figure 1.2. Daily otolith growth patterns from Atlantic cod (*Gadus morhua*). Each dark ring corresponds to 1 day of growth. Scale bar = 10 μm (from Campana and Thorrold 2001).

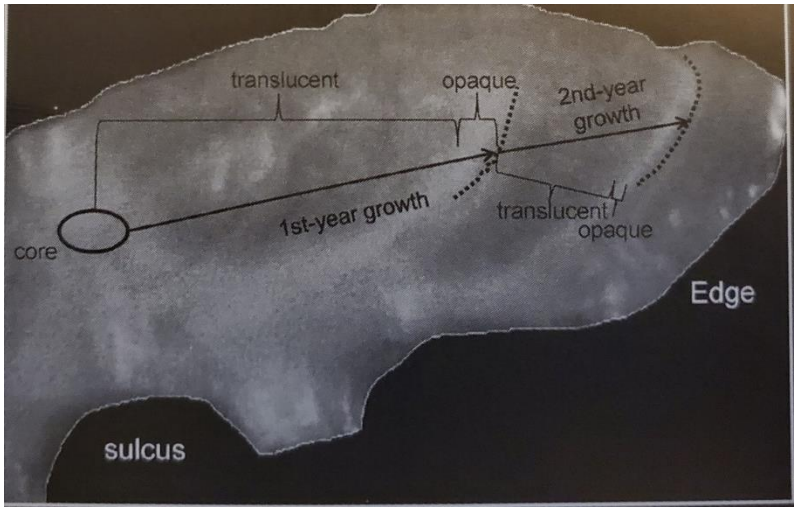


Figure 1.3. Annual growth patterns on a sectioned Largemouth bass (*Micropterus salmoides*) sagitta, examined under reflected light (from Long and Grabowsky 2017).

Numerous studies have been done examining daily growth ring formation since their discovery in the 1970's (Pannella 1971; e.g. Jones and Brothers 1987). While many of these have found that daily ring formation is regular enough to reliably determine age, numerous factors can weaken this ring-age relationship. Feeding regime was investigated by Jones and Brothers (1987) in striped bass and found to significantly alter the regularity and timing of ring deposition. Temperature has also been shown to have an impact on daily increment formation (Bestgen and Bundy 1998; Song et al. 2009; Long and Porta 2019), though this appears to be less important than for annual growth bands (Campana 2001). Photoperiod can also have an impact in increment formation, though it is possible that this effect is more related to feeding success than direct light availability (Morales-Nin 2000). This method begins to fail in some species when applied to fish greater than 100 days old due to difficulty of ring differentiation (DiCenzo and Bettoli 1995), though use of electron microscopy is able to extend this period significantly (Waldron and Kerstan 2001).

Annual increment formation has been well established in numerous taxa throughout all life stages (Campana and Neilson 1985). The time of annulus formation varies by location, seasonality, and age (Buckmeier et al. 2017). The important point of formation time, however, is that it does represent annual periodicity. Whether the light or dark bands are counted as "annuli" varies by taxa, with some degree of confusion across the literature as to which should be used, but as long as it is consistently applied either is considered acceptable (Wilson et al. 1987; Quist et al. 2012; Long and Grabowski 2017). Stress and change of environmental conditions can result in the formation of a false ring, but this is highly dependent on the population under examination (which highlights the need for population specific validation, discussed below) (Long and Grabowski 2017).

Otolith preparation can vary depending on the taxa and age of the fish (Campana 2001). The alternating bands can either be read directly from the surface of the otolith, transverse breaking of the otolith followed by burning or baking it, or thin sectioning, mounting, and polishing. Surface reading of unprocessed otoliths is the primary method of ageing for daily growth, as the size of otoliths from young fish might prohibit further manipulation (Long and Grabowski 2017). While surface reading is the easiest and most common method, requiring the least amount of preparation, it is associated in some species with a significant underestimation of age (Beamish and McFarlane 1995; Long and Grabowski 2017). This is primarily due to the visual underexposure of annuli, which causes progressively more error as new growth is added on. The non-isometric growth which many fish species experience as they reach higher age groups compounds this problem (Beamish and McFarlane 1995). Transverse breakage and burning of the otolith can help compensate for this error by exposing hidden annuli, but this method requires

more sample preparation in exchange. Thin sectioning likewise is able to expose more hidden annuli, but long-term storage of mounted sections can cause yellowing or cracking, in addition to the added labor of the sectioning and mounting itself (Beamish and McFarlane 1995; Campana 2001; Long and Grabowski 2017). Preparation still varies dramatically across studies, though some authors push for more uniform acceptance of sectioned otoliths as the most reliable option (Winkler et al. 2019).

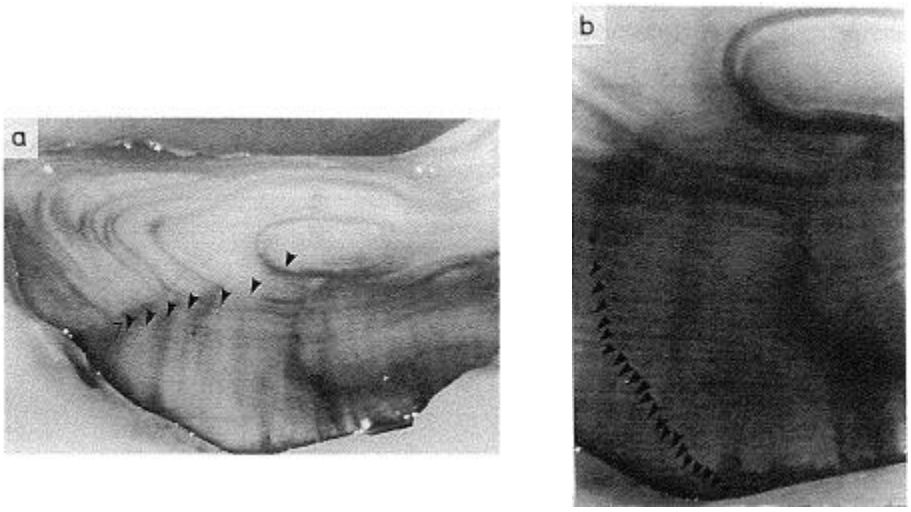


Figure 1.4. Sectioned walleye pollock (*Gadus chalcogrammus*) otoliths, with an estimated age of 28 years. The surface predicted age for the same individual was 20 years, indicating that age underestimation might be significant for surface reading from higher age classes (from Beamish and MacFarlane, 1995).

Vertebrae

The primary use of vertebrae for ageing purposes is in taxa where a relative lack of calcified structures prevents most other ageing methods (such as elasmobranchs). While vertebrae are sometimes used in ageing bony fishes (McCarthy and Minckley 1987), focus will primarily be paid here to their usage in ageing cartilaginous fishes. The beginning of serious elasmobranch age examination began significantly later than was

seen in teleost fishes (Calliet et al. 1986; Prince and Pulos 1983), which has resulted in a degree of lag in validation of taxa (Harry 2018).

The terminology used for vertebral ageing varies slightly from that of otolith ageing, but there is some overlap. Similar to otoliths, vertebral ageing relies upon alternating "opaque" and "translucent" bands which spread radially and longitudinally from the center of the vertebrae. Each "band pair" consisting of an opaque and translucent band comprises an annuli (Cailliet et al. 2006). Daily ring formation has not been found in vertebrae. Vertebral bands also differ from otoliths in chemical composition: elasmobranch vertebrae consist of different combinations of hydroxyapatite $3(Ca_3PO_4)_2$ and organic matrix, with the different ratios of these constituent parts resulting in the lighter and darker bands observed (Kerr and Campana 2014).

Preparation of vertebrae for ageing generally involves sectioning. Some studies do make use of whole vertebrae, but this is largely cautioned against as it can cause significant age underestimation (Bennett et al. 1982; Kusher et al. 1992; Dwyer et al. 2016; Vinyard et al. 2019). More commonly, vertebrae are cut longitudinally using a jeweler's saw or similar tool to create thin sections, either with or without embedding in resin (Humason 1972; Smith 1984; Kusher et al. 1992). Staining is also often employed to help visually distinguish band pairs, usually via a silver nitrate solution which darkens the "opaque" bands (Cailliet et al. 1983).

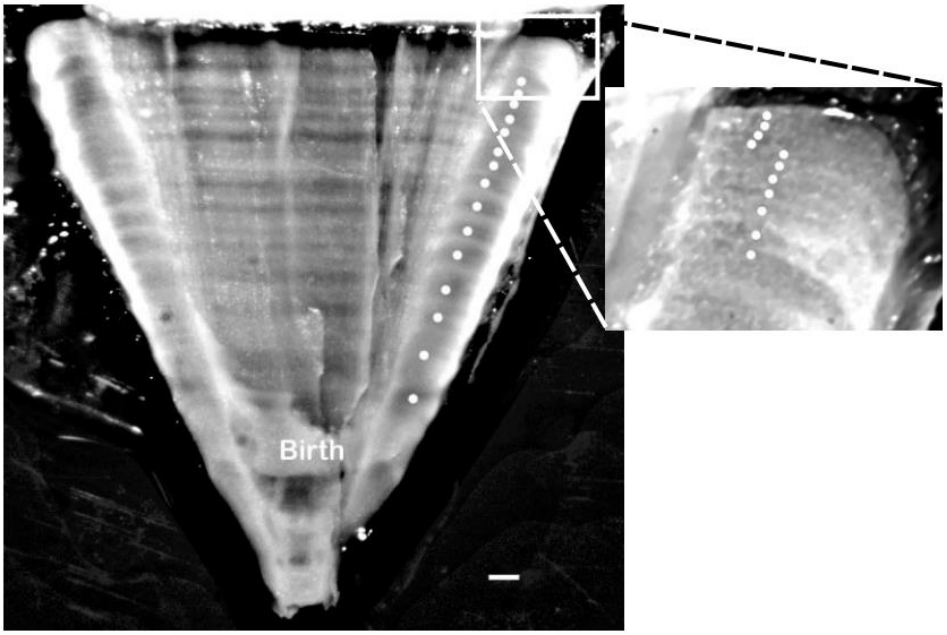


Figure 1.5. Thin vertebral section from a porbeagle shark (*Lamna nasus*). Each dot marks a suspected annulus, with the zoomed section highlighting the tight grouping seen as a fish ages. Scale bar = 1mm (from Campana et al. 2002).

Periodicity of band pair formation can vary by life stage within a species, which requires attempts to validate vertebral ageing be highly age and taxa specific (Natanson et al. 2002; Campana 2001). There is significant evidence that vertebral ageing has resulted in systemic underestimation of age in multiple shark taxa; a recent review suggested that of 29 genera investigated, 9 were likely underaged, comprising approximately 30% of the populations examined (this percentage goes up slightly when examining only studies which utilized vertebrae, though these studies did comprise 89% of the studies used) (Harry 2018). A review of studies which utilized vertebral ageing for sharks found that the average percent error (discussed in more detail below) was frequently twice as large as was commonly reported for teleost otoliths (Campana 2001). This error is largely focused on older elasmobranchs, though there is some debate as to its primary source. Some suggest that the primary issue is the crowding which results from slower growth

later in life, which renders the bands indistinguishable (Francis et al. 2007; Chin et al. 2013). However, it has been pointed out that this suggests higher resolution would eventually counteract this issue, and that numerous studies examining the use of radiography, x-ray fluorescence microscopy, and computerized tomography would have been expected to do so (Harry 2018). It seems more likely that this underageing trend is the result of growth bands either ceasing to form or changing periodicity as fish age (Francis et al. 2007; Natanson et al. 2016; Kinney et al. 2016), or some degree of combination between the two factors (Harry 2018).

Fin Rays and Spines

Fin rays and spines are also often used for ageing, sometimes in taxa in which other methods have been unsuccessful, taxa for which nonlethal sampling is strongly preferable, or as a comparison structure for other ageing methods (Beamish and McFarlane 1985; Buckmeier et al. 2002; Campana et al. 2006; Rude et al. 2013). The ageing process and structure of rays and spines are quite similar to that of other deposition ageing methods. Chondrichthyan dorsal fin spines, when present, consist primarily of enamel-coated dentine, similar to placoid scales (Whitledge 2017). Teleost fin spines consist of a central lumen, from which growth radiates outwards (Whitledge 2017). The result of these differing compositions is that reading chondrichthyan spines is fundamentally different than reading teleost spines- for the former, spines are read by band accumulation distal to the notochord, while the latter is read in transverse sections with bands radiating outwards (Campana et al. 2006; Kopf and Davie 2011; Whitledge 2017; Fischer and Kooch 2017).

Fin rays and spines are prepared in a manner analogous to that of vertebrae, with thin cross sections obtained with a low speed saw or similar tool (Witt 1961). Most modern studies encase rays and spines in epoxy before this sectioning occurs, to prevent distortion caused by cutting, to ease in the process of cutting itself, and to prevent breakage of brittle structures (Koch and Quist 2007). Age is then read, usually with a compound or dissecting microscope (Fischer and Koch 2017).

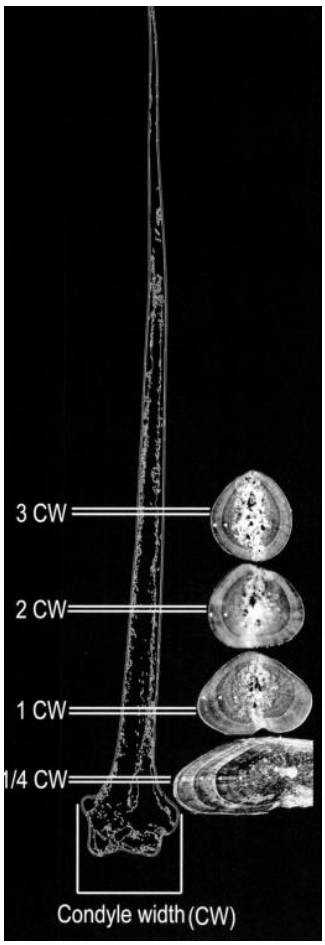


Figure 1.6. Diagram showing the location of transverse cuts used to section fin spines in striped marlin (*Kajikia audax*). The authors found that the location of the section along the shaft and the location of the removed spine had significant effects on age determinations (from Kopf and Davie 2011).

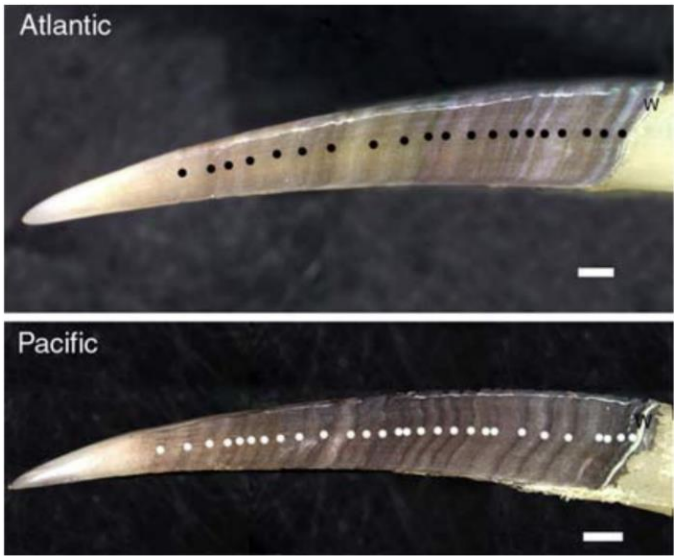


Figure 1.7. Dorsal spines from spiny dogfish (*Squalus acanthias*) collected in the Atlantic (top) and Pacific (bottom) oceans. Dots represent annual growth bands for each (from Campana et al. 2006).

The accuracy of fin ray and spine ages are often comparable to that produced by otoliths (e.g. Quist et al. 2007) (Fischer and Koch 2017). Fin ray and spine derived ages are often greater than those produced by scales and vertebrae (e.g. Phelps et al. 2007), though the systemic underageing of those structures previously mentioned means that this might represent a higher degree of accuracy. Past a certain age, however, spine and ray ages often fall behind those estimated by otoliths (Welch et al. 1993; Phelps et al. 2007). Despite their general similarity to otolith produced ages, however, studies have also found that fin spines and rays produce ages that are more variable (Maraldo and MacCrimmon 1979) and less precise (Isermann et al. 2003) than those produced by otoliths. Some of this error is likely due to the effect vascularization has in reducing visibility of early annuli (Kopf and Davie 2011), but the fact that fin rays can be regenerated if damaged or lost also introduces some degree of inaccuracy in ages derived from these structures (Johnson and Weston 1995; Witten and Huysseune 2009).

Validation

Any discussion on the accuracy of ageing techniques necessarily implies that the actual age of a specimen is known. As these hard-structure ageing methods are the ones most commonly used to determine age, a method of externally validating the assumptions behind them (such as the annual creation of a new annulus) is necessary. This is done through validation studies, wherein the growth of ageing structures is tested in fish of verified age (though it should be noted that usage of "validation" can vary slightly, with some authors meaning validating the ages assigned to individuals and some meaning validating the depositional patterns which are used) (Beamish and McFarlane 1983; Buckmeier et al. 2017). Methods of validation vary, but are usually more expensive or less practical in wide application than traditional hard structure ageing (hence its popularity and widespread use). Validation methods include using known-age fish, mark-recapture, bomb radiocarbon analysis, radiometric analysis, and marginal increment/ edge analysis (Campana 2001).

Validation is an absolutely vital part of any ageing technique, but also an easy one to get wrong (Beamish and McFarlane 1983; Campana 2001). Numerous examples exist of researchers using a structure which has been validated for the species in question yet doing so incorrectly (e.g. Steffenson 1980; Pratt and Casey 1983). There are multiple reasons for this, but they often involve not validating the age group under examination (Campana 2001). Growth increments of immature fish are rarely created with the same periodicity of growth increments created after maturity, as has been discussed above in a structure-specific manner (Casselman 1987; Natanson et al. 2002; Campana 2001). The result is that studies which are operating with a "validated structure" for a specific taxa

might still result in wildly incorrect age determinations if the validation itself is not specific enough for the population being examined.

The ideal situation is often utilizing fish of absolutely known age (such as those captively hatched) to validate the periodicity of ageing structure growth and true age simultaneously (Campana 2001). By gradually collecting individuals over a long period of time, the development of key structures can be documented and tested for regularity across ages. This is, however, impractical in many cases. Marine species often require conditions which are difficult to match, and many are long-lived enough to make a study of this kind infeasible (Buckmeier et al. 2017). The specific growing conditions of a population, including feeding regime and temperature patterns, can also significantly affect the development of hard structures, which makes many captive-reared populations inappropriate models for wild populations (Buckmeier et al. 2017). This method is quite popular in validating daily increment growth in otoliths, however, due to the more limited nature of the experiment duration required (Geffen 1992; Jones and Brothers 1987; Campana 2001). Studies utilizing captive-hatched but wild stocked populations are also able to avoid the issues of different development conditions affecting structure growth, but are only practical for a limited number of situations where native recruitment is not possible or where the native population is able to be distinguished (Buckmeier et al. 2017). Alternatively, stocked fish chemically marked before release can provide the same degree of accuracy across more disparate taxa, though still with a likely limit of study length (Campana 2001; Buckmeier et al. 2017).

Mark-recapture age validation experiments involve catching a specimen, chemically marking it (often with oxytetracycline (OTC)), and then releasing it back into

its environment. For individuals which can be visually aged (such as a neonate), this can also be done via physical tagging. A resampling is done after the desired period. The OTC markers incorporate into calcified structures and indicate the age at the time of marking via fluorescence, and as the time after that is definitively known, the assumptions behind the frequency of band formation can be tested (Holden and Vince 1973; Campana 2001; Schill et al. 2010). Similar to known-age validations, however, the amount of time which can reasonably be covered by this method is generally quite low; often, longer times at liberty reduce the number of individuals which are successfully recaptured (Beamish and McFarlane 2000; Natanson et al. 2002). Some studies avoid this problem by tagging wild-caught fish and then rearing them in enclosures, but this reintroduces the problems of mimicking a natural environment which were discussed in the previous paragraph (Schmitt 1984; Campana 2001).

Bomb radiocarbon analysis relies on the different amount of ^{14}C present in individuals born between 1958-1968 as a result of nuclear testing. Prior to 1958 relatively little ^{14}C was present in the atmosphere, but this increased dramatically over the following decade. This essentially works as a chemical marking placed upon all fish during these years, with the concentration of ^{14}C working as a comparison chronology (Campana 2001). The result is an ageing technique which has resolution between 2-5 years, depending on the target population, and can work with any individual with a hatch date throughout the 1960's (Campana 2001; Buckmeier et al. 2017). Bomb radiocarbon ageing has been particularly important in the validation of elasmobranch ageing structures (Francis et al. 2007; Andrews et al. 2011; Harry 2018). Care must be taken to develop an appropriate reference chronology, given the difference in ^{14}C between

freshwater and marine environments, and creation of an outside chronology is not always feasible (Campana and Jones 1998; Campana 2001). This approach has been less applicable for freshwater species, which are often less long-lived than marine taxa, but it has been successfully applied for some (Bruch et al. 2009). The usefulness of this technique has continued to decline as time since the 1960's increases and is expected to continue to do so (Buckmeier et al. 2017).

Radiometric/ radiochemical dating relies instead on the decay of naturally occurring isotopes. Once the core of the otolith is fully formed, no new material is added except on the outer surface. This means that isotopes which have been incorporated into it are fixed and therefore able to be aged via their respective half-lives. This method is largely accurate, though only with a resolution power of roughly 5 years it is mostly applicable in longer-lived species (Campana 2001). Its utility is also wholly reserved for bony fishes, as other structures (scales, spines, bones, etc.) continue to grow and sometimes reabsorb, which varies the base amount of each isotope present over time (Campana 2001; Buckmeier et al. 2017).

Edge analysis and marginal increment analysis (EA and MIA, respectively) are techniques to validate periodicity of growth structure creation without validating absolute age (Campana 2001; Buckmeier et al. 2017). Edge analysis involves recording the opacity of a structure's edge at various points throughout the year (Holden and Vince 1973; Casselman 1987; Lapropoulou and Papaconstantinou 2000). By recording this information regularly (such as monthly), the creation of an annual cycle consisting of a dark band and a light band can theoretically be verified. Marginal increment analysis is similar to this, but rather than a binary distinction of edge state it is recorded as the

distance between the last annulus and the edge, either in length or as a proportion of annulus completion (Compana 2001; Buckmeier et al. 2017).

Together, MIA and EA represent the majority of validation studies done (Campana 2001; Buckmeier et al. 2017). Despite this, it is cautioned against by Campana (2001) as the "most commonly used, and the most likely to be abused" of the validation techniques covered in his review. There are many reasons for this. For one, the state of a given edge can be difficult to discern and is open to subjective interpretation. This is described by Campana as a commonly used "'looks like a cycle to me'" approach. Edge analyses also are easily misapplied, as previous validations have been used on groups older than those under examination; significant evidence suggests that the timing and periodicity of edge state vary with age (Campana 1984; Campana 2001; Buckmeier et al. 2017). Marginal increment analysis is often more statistically rigorous, and is the most common validation approach used in chondrichthyan age studies (Calliet and Goldman 2004; Calliet et al. 2006). However, MIA and EA represent the cheapest and easiest option for validation. No captive rearing, chemical marking, or recapturing is required-fish are simply sampled throughout the course of a year (Buckmeier et al. 2017).

Errors in Ageing

Before discussing error, attention should be drawn to the difference in ageing studies between precision and accuracy. Accuracy reflects how well a determined age matches the actual age, and is largely determined through validation studies. Precision denotes the repeatability of the age determination itself (Kalish et al. 1995; Campana 2001). While high precision is important, it is not itself an indication of accuracy.

Estimated ages which are significantly different than actual age are very capable of having a high degree of precision (Campana et al. 1990; Campana 2001). There is some degree of misuse in the literature between these terms, but standardization of ageing terminology has long been pushed for in the field (Beamish and MacFarlane 1985; Campana 2001; Panfili et al. 2002; Buckmeier et al. 2017). When in doubt, we have gone with the definitions used by Campana (2001), as this is generally the most authoritative and oft-cited review of ageing error.

Fish ageing has two well established sources of error. The first source of error can be defined as process error. This is the result of the structure not growing at the expected rate, which in turn results in a greater or (more commonly) fewer number of annuli than should be present. Thorough validation studies help to minimize this, but with so many factors impacting structure growth it is difficult to validate each population to such a degree that the risk of this error is fully dismissed. The second source of error is in the process of age determination itself (Campana 2001). This can be due to issues of sample preparation, reading, or interpretation (Beamish and McFarlane 1995). Attempts to minimize this primarily involve changing the process of reading itself, often by limiting the specimen information available to readers, utilizing multiple readers to establish consensus, having each reader examine each structure multiple times, and general quality assurance/ quality control procedures (Campana 2001; Calliet et al. 2006; Buckmeier et al. 2017).

Reporting Accuracy and Precision

The way in which both accuracy and precision are reported has varied widely (Beamish and Fournier 1981; Change 1982; Campana et al. 1995; Campana 2001). In regards to precision, percent agreement was the most commonly used measure prior to the 1980's when alternatives were proposed, and despite its oft-commented on inadequacies (Campana 2001) it remained the most highly used according to a 2005 review (Morison et al. 2005). Percent error neither accounts for true age of fish or the degree of error found, as it simply records the percent of age counts which agree with one another (either by the same or multiple readers) (Buckmeier et al. 2017). The result is that a 1 year ageing error on a 2 year old fish is reported the same as if on a 50 year old fish, and a 5 year ageing error on a 2 year old fish is reported the same as the 1 year error. This issue prompted Beamish and Fournier (1981) to propose the use of average percent error (APE, referred to as the index average percent error (IAPE) when averaged across multiple fish). This was defined by them as:

$$\mathrm{APE_j} = 100\% \times \frac{1}{R} \sum_{i=1}^{R} \frac{|x_{ij} - x_j|}{X_j}$$

where X_{ij} is the *i*th age estimate for the *j*th fish, X_j is the average age estimate for the *j*th fish, and R is the number of times each fish was aged. This allowed the magnitude of error to be weighted by the presumed age of the fish, and allowed comparison between the precision of separate readers (Beamish and Fournier 1981).

Shortly after this, Change (1982) proposed a modification of this equation, using the standard deviation of the average determined age rather than the absolute deviation. He defined this new equation as:

$$CV_j = 100\% \times \frac{\sqrt{\sum_{i=1}^{R} \frac{(X_{ij} - X_j)^2}{R-1}}}{X_j}$$

where CV_j is the coefficient of variation for the *j*th fish and the rest of the variables are comparable to those of Beamish and Fournier's (1981) APE. These measures are very similar; Kimura and Anderl (2005) found that for most samples CV is equivalent to $\sqrt{2}\ x\ APE$. Given their close relationship, Campana (2001) found no clear preference of one over the other, but did point out that CV is more statistically rigorous. His review of 131 ageing studies found that 57% used CV, though APE and CV each comprised approximately 50% of papers examining only annual ageing. These are now the standard suggested measures of precision, though with some exceptions. For instance, a study calculating the strength of a given year class might find that error is more helpfully expressed as an absolute, rather than weighted by age (Buckmeier et al. 2017).

Accuracy reporting is slightly more standardized, though still with some variations. A review by Campana et al. (1995) examined percent agreement with actual age, age difference plots, parametric t-tests, and nonparametric Wilcoxon matched-pairs rank tests. The result of this was the usage proposal of an age bias plot, where estimated age from one reader is plotted against estimated age of another reviewer by age classes, with estimated age expressed as both the mean and a 95% confidence interval. Campana

et al. (1995) found that this method was better at drawing attention to both linear and nonlinear ageing biases. By substituting actual age for reader age on one axis, the accuracy and bias in estimation are readily apparent (Campana 2001; Morison et al. 2005; Buckmeier et al. 2017). In line with the continued utility of age bias plots, an R library called “FSA” has a built in ageBias plot function (figure 1.9, below).

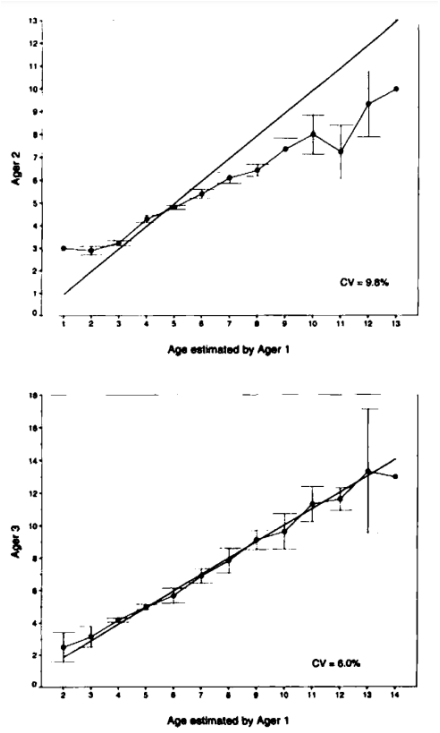


Figure 1.8. Age bias plots from 2 readers, with the point representing the mean age determination and the bars representing 95% confidence intervals. Both linear bias (such as that seen in the top graph) and nonlinear bias are able to be detected using this method (from Campana et al. 1995).

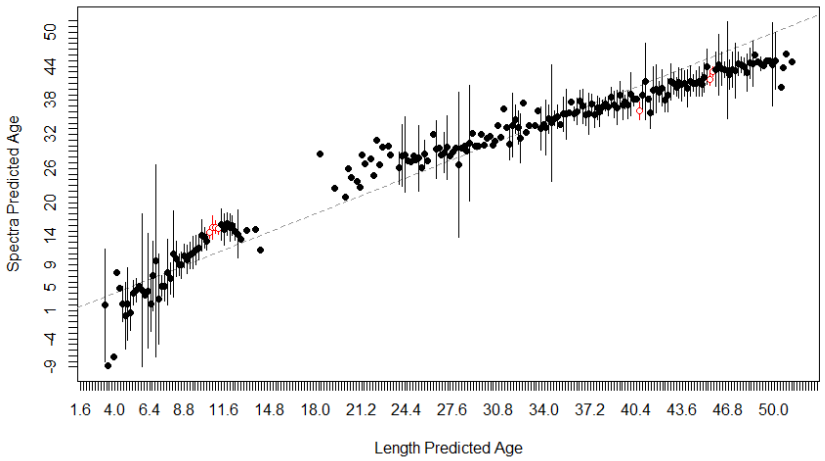


Figure 1.9. Age bias plot comparing predicted ages from near infrared spectroscopy (NIRS) and length-predicted ages of a young group of striped bass (*M. saxatilis*) (from chapter 2 research).

Current recommendations (and those since the mid 1990's) call for the reporting of age bias plots for measuring accuracy and CV for measuring precision (though, as mentioned above, some authors do not strongly favor CV over APE) (Campana et al. 1995; Campana 2001; Morison et al. 2005; Buckmeier et al. 2017). However, the methods chosen will always depend on the purpose of the experiment (Buckmeier et al. 2017). Measures of both accuracy and precision utilized will therefore continue to vary with experimental design, but an increasing avoidance of simple percent agreement has significantly helped with efforts to compare results between studies. In the following chapters, most results will be expressed in terms of APE and measures of bias, in addition to a number of NIRS-specific metrics.

FT-NIRS ageing will be introduced more thoroughly in the next chapter. The following chapter will also investigate the use of known age striped bass for FT-NIRS ageing analysis, and examine the importance of storage method on scan and model

accuracy. It will also investigate the use of whole fish, which is untested in FT-NIRS ageing but, I will argue, is based on a sound theoretical backing. In the third chapter I will test the utility of FT-NIRS ageing on a species of elasmobranch which is of significant commercial interest, and will examine sample preparation techniques as they relate to FT-NIRS scans. Finally, in the fourth chapter I will utilize FT-NIRS ageing for a species of skate using embedded and mounted vertebral sections. As these sections have been traditionally aged previously, I will create models to test the fidelity of spectroscopic ageing to originally assigned ages. The usefulness of epoxy embedded and mounted samples for FT-NIRS fish ageing has not been investigated, but as many historical and reference samples are thus prepared, this knowledge would serve to expand the potential calibration samples available. The overall goal of each of these chapters collectively is to expand the knowledge available about the practicalities of broader incorporation of FT-NIRS in fisheries science.

CHAPTER 2

FT-NIRS AGE RESOLUTION DETERMINATION USING WHOLE STRIPED BASS LARVAE AND JUVENILES OF KNOWN AGE

Introduction

Researchers have begun to explore the use of Fourier-Transform Near Infrared Spectroscopy (or FT-NIRS) as a tool to age fish more quickly and easily (Wedding et al. 2014; Rigby et al. 2014; Robins et al. 2015; Rigby et al. 2015; Helser et al. 2019; Passerotti et al. 2020a). FT-NIRS is a chemometric analysis technology which has found varied applications in pharmacology (Roggo et al. 2007), biomedical research (Macnab 2009) and agriculture (Solberg et al. 2003; Bobelyn et al. 2010; Wedding et al. 2009). In ecological fields, NIRS application has included beetle species identification (Teixeira et al. 2015), frog sex discrimination (Vance et al. 2015), ostrich dietary nutrition (Landau et al. 2006), and much more (Vance et al. 2016). FT-NIRS works by irradiating a sample with near infrared energy (12,800 – 4000 cm^{-1} wavenumbers). Energy is absorbed by the sample corresponding to the vibrational frequency of bonds within the molecules present. The energy absorbed, and the light which is reflected, is recorded in a spectrum, with peaks representing the vibrational frequency of bonds present in the sample. This technology allows for fine-scale discrimination between samples of differing chemical composition, though without definitively identifying the molecular structures responsible for this differentiation (Miller 2001).

Table 2.1. Previously reported calibration model results for FT-NIRS ageing of fish. Many of the studies listed included additional components not included here, such as testing samples gathered at separate times (e.g., Wedding et al. 2014), different locations (e.g., Robins et al. 2015), or both (e.g., Helser et al. 2019). In these cases, only the combined model results are presented. When possible, the largest or most representative model was chosen for inclusion. R^2 = coefficient of determination; RMSE = Root Mean Square Error; % RMSE = RMSE/ maximum age of sample set * 100. RMSE is given in years, unless denoted by *, which indicates RMSE is in days.

Species	Structure	n	R^2	RMSE	%RMSE	Study
Squalus megalops	Dorsal fin spine	97	0.82	2.41	9.64	Rigby et al. 2014
Squalus montalbani	Dorsal fin spine	95	0.73	2.96	9.54	Rigby et al. 2014
Squalus megalops	Vertebrae	97	0.89	1.85	7.4	Rigby et al. 2014
Squalus megalops	Fin clips	97	0.76	2.67	10.68	Rigby et al. 2014
Lutjanus malabaricus	Otoliths	169	0.93	1.35	5.87	Wedding et al. 2014
Gadus chalcogrammus	Otoliths	202	0.95	0.78	4.87	Helser et al. 2019
Lates calcarifer	Otoliths	298	0.86	0.75	6.25	Robins et al. 2015
Pagrus auratus	Otoliths	306	0.88	1.53	6.12	Robins et al. 2015
Lutjanus campechanus	Otoliths	108	0.91	6.33*	5.28	Passerrotti et al. 2020a
Lutjanus campechanus	Otoliths	508	0.94	1.54	4.97	Passerotti et al. 2020b
Sphyrna mokarran	Vertebrae	76	0.89	0.87	8.52	Rigby et al. 2015
Carcharhinus sorrah	Vertebrae	99	0.84	0.88	8.97	Rigby et al. 2015
Lutjanus campechanus	Otoliths	333	0.94	1.33	4.43	Barnett et al. 2019
Raja rhina	Vertebrae	324	0.88	1.41	7.83	Arrington et al. 2019
Oncorhynchus tshawytscha	Otoliths	245	0.81	0.36	6	Claiborne et al. 2019

In fish, the technique has found success in otoliths (Wedding et al. 2014; Robins et al. 2015; Helser et al. 2019; Passerotti et al. 2020a), chondrichthyan vertebrae (Rigby

et al. 2014; Rigby et al. 2015), chondrichthyan dorsal fin spines (Rigby et al. 2014), and chondrichthyan fin clips (Rigby et al. 2014). Different mechanisms have been proposed for the FT-NIRS – age relationship, including protein composition in the otolith organic matrix (Helser et al. 2019), and the deposition of hydroxyapatite $3(Ca_3PO_4)_2 \cdot Ca(OH)_2$ in chondrichthyan fin clips, vertebrae, and dorsal fin spines (Rigby et al. 2014). No definitive identity for the responsible molecule (or molecules) has been determined, however.

It is partly this uncertainty in causal molecules which prompted the present exploration. As the mineralization of fish scales is a continuous process (Schönbörner et al. 1979), it was hypothesized that molecular differences on the exterior surface of teleost fishes, detectable with FT-NIRS, would closely correspond to age during early life stages. In order to also eliminate any uncertainty associated with traditionally determined ages, a hatchery produced species was chosen. This choice allows for a more standardized sample age distribution than has been used previously in FT-NIRS ageing investigations, which, combined with possessing known-age validation and calibration samples, will clarify potential limits of FT-NIRS ageing in daily resolution. While recent studies have begun examining the utility of FT-NIRS ageing on a daily scale, they have relied upon traditional ageing methods for model creation and testing, compounding any potential errors therein (Passerotti et al. 2020a; Helser et al. 2019).

Accuracy comparisons between FT-NIRS ageing studies and traditional ageing studies are frustrated by multiple factors. First, many ageing studies report results as percent agreement, defined as the percent of samples which were assigned to their correct age group (Branigan et al. 2019; Copeland et al. 2007; Klein et al. 2017; Secor et al.

1995). As FT-NIRS regression models predict ages as a continuous variable rather than a discrete one, this method would require modification in order to apply to studies analyzing ageing accuracy of FT-NIRS. As the percent agreement method also fails to indicate significant portions of relevant data (Campana 2001; Beamish and Fournier 1981), its usefulness as a measure of FT-NIRS ageing accuracy is doubtful. Use of the average percent error (APE) equation suggested by Beamish and Fournier (1981) could, with minor modifications to remove inclusion of among-reader consensus, allow for stronger direct accuracy comparisons. Second, directional bias in prediction accuracy is quite common in traditional ageing studies (e.g. Bruch et al. 2009; Long and Porta 2019; Song et al. 2018), while most FT-NIRS studies have not shown a predictable cross-taxa bias (Wedding et al. 2014; Rigby et al. 2014; Rigby et al. 2015; Passerotti et al. 2020a; Claiborne et al. 2019; Arrington et al. 2019; Helser 2019). Hard structure ageing requires consistent (or at least predictable) structure growth, which can vary widely based on numerous environmental factors, such as temperature (e.g. Song et al. 2009), feeding (e.g. Jones and Brothers 1987; Bestgen and Bundy 1998), and latitude (e.g. Albuquerque et al 2019). The overall result is a tendency to underage older fish. The lack of consistent bias suggests that, while FT-NIRS predicted ages might result in a less skewed view of population age structures, error in prediction might be more difficult to compensate for in FT-NIRS predicted ages.

Striped Bass (*Morone saxatilis*) is one of the most important commercial fishery species along the Atlantic United States Coast (Liao et al. 2013). In South Carolina, little natural recruitment takes place due to spawning condition requirements. Most striped bass populations in the state are therefore maintained by the stocking of hatchery-reared

fish (Wildlife and Freshwater Fisheries Division 2009). As a commonly available hatchery species which serves an important economic function, it was deemed a strong choice for the present study objectives. These objectives included 1) determining whether whole-fish FT-NIRS scans would correlate with age as strongly as is seen using otoliths, 2) testing the feasibility of FT-NIRS to distinguish daily ages of a commercially important stock, 3) validating the utility of FT-NIRS ageing models in a setting without error brought in by use of traditionally assigned validation and calibration ages, 4) simulating differences in model accuracy which could have occurred if validation and test ages had been determined through increment counts, 5) determining the minimum number of calibration samples needed to produce a strong predictive model, and 6) comparing the comparative utility of FT-NIRS ageing to an age-length regression age prediction.

Methods

Stock

During 2019, and for days 1-20, samples were taken from the Jack Bayless Fish Hatchery (St Stephen, SC) managed by the South Carolina Department of Natural Resources (SCDNR). The broodstock (1 female, 3 males) used to produce these fry were collected from the Santee River/ Rediversion Canal. The fish used for this study were all part of the same family group to limit the possible influence of genetic differences in growth. Once hatched, fry were kept in a 4-foot circular holding tank. No additional food was provided for the first 5 days after hatching, as the fry lived on their attached yolk sacs. After 5 days, a diet of *Artemia* (brine shrimp) was provided twice daily.

After 15 days at Jack Bayless Fish Hatchery, a subset of the Bayless stock was transferred to the Cohen Campbell Fisheries Center (West Columbia, SC) managed by the SCDNR. The fish to be transported were packaged in plastic bags filled with oxygen and sealed in Styrofoam coolers, which were then taken by truck to Cohen Campbell. Once at the Cohen Campbell facility, fish were acclimated to the ponds for approximately 45 minutes before stocking, or until the pond temperature was reached in the bags. All samples used from this location were stocked in 3 half-acre ponds. Fish at Cohen Campbell Fisheries Center were not fed artificially but grazed, *ad libitum*, on naturally occurring invertebrates (e.g., *Daphnia*, Copepoda, Rotifera, etc.) that colonize each pond. Invertebrate growth and population density were encouraged via cotton seed meal.

Sample Collection

Day 1 specimens were collected roughly 6 hours after hatching, with each subsequent day (through day 20) being sampled approximately 24 hours later. After a subset of fish were transferred to Cohen Campbell Fisheries Center at day 15, sampling continued on the stock still at Jack Bayless Fish Hatchery until day 20. Sampling continued at Cohen Campbell from days 24-49, with all specimens collected at night. A weak light attraction response, as well as small size relative to the new ponds, prevented samples from being taken at Cohen Campbell from days 21-23. Sampling at the Cohen Campbell facility occurred between 22:00 and 04:00 EST, using a large light to attract fish and a fine-mesh net to capture individuals as they approached the surface. Sampling continued each night until approximately 20 individuals were captured per storage method. As the Cohen Campbell individuals were sampled approximately 12 hours later

than those sampled from days 1-20, they were assigned ages in half days (e.g., samples collected during the night 30 days after hatching were assigned an age of 30.5 days).

I set, *a priori*, a target number of 40 individuals per sampling event, with 20 stored in 95% EtOH and 20 stored in water to be frozen. This was either achieved or nearly achieved (n = 18+) in days 1-9, 12-20, 24-29, and 31-49 for the EtOH stored set. No samples were obtained for days 10 or 11, or for days 21-23, as explained previously. Only 10 samples were obtained for day 30, all of which were included in the subsequent analyses. Samples were immediately stored in 95% EtOH until further analysis. Frozen samples achieved or nearly achieved (n = 19+) the desired sample size for all days except 21-23 and 30. Frozen samples were stored in water at -40º C until testing.

Data Collection

The standard length of each EtOH preserved specimen was taken from the left side of each individual by laying them flat on a ruler's surface. For days 1-20, and prior to strong differentiation of the caudal fin, tissue opacity was used to determine the posterior end of each length measurement. Measuring was done under a dissecting microscope for days 1-20, and visually for days 24-49. In both cases, samples were laid on a ruler, with results recorded to the nearest tenth of a millimeter. To prevent abnormalities in exposed tissue from confounding the results, any visually damaged samples (n = 11) were removed from further analysis.

NIRS scanning was done at $16cm^{-1}$ resolution, with a repetition of 64 scans per sample. This was performed with a Bruker Matrix-I Near Infrared Spectrometer with a 22-mm diameter sample window using OPUS software (version 8.2; Bruker Scientific,

Billerica, MA). Scans used the entire fish, positioned with the right operculum in the center of the aperture. Larger (day 24-49) samples were air dried following length measurement and prior to scanning in order to prevent EtOH pooling on the aperture. This was not necessary for smaller, day 1-20 fish, which began visibly shriveling less than 3 minutes after being removed from EtOH filled storage tubes, and which dried sufficiently to prevent pooling after less than a minute. Frozen samples were allowed to thaw before being gently patted dry to prevent water pooling on the window. A 19-mm gold transflectance stamp was placed above each sample before scanning in order to standardize the NIR light lengths obtained. Scans included the entire NIR spectrum (3600 - 12,000 cm^{-1}). The 64 scans of each sample (n = 991 for EtOH set and n = 1115 for frozen set) were averaged together to create a single spectrograph for each sample, for each storage method.

Data Analyses

All analyses regarding spectral data were done utilizing OPUS software (version 8.2; Bruker Scientific, Billerica, MA). Unless otherwise specified, all statistical analysis and simulations were done using Microsoft Excel 2016. Multivariate data from each of the spectrographs of the remaining samples (n = 991 EtOH and n = 1115 frozen) were modeled relative to their known ages using a partial least squares regression (PLSR) (Chen and Wang 2001). Spectral data was then tested to determine appropriate pretreatments, and associated wavenumber ranges for each pretreatment, with the goal of minimizing RMSE while avoiding utilization of spectral noise, which could result in overfitting.

All spectrographs were utilized in a leave-one-out analysis comparing age and spectra for each storage set. In this analysis, a regression line is created utilizing all spectrographs but 1, which is then plotted along the regression line created. This is repeated, with each sample being excluded in turn. The results are reported in R^2 (coefficient of determination), RMSECV (root mean square error of cross validation), and RPD (residual prediction deviation).

Following this, samples were split by storage method into calibration and test sets. To test the importance of calibration set size on model accuracy, 5 analyses were performed at different calibration set frequencies (0.5, 0.33, 0.25, 0.2, and 0.1 of all samples, corresponding to calibration set sizes of $n = 495$, $n = 330$, $n = 247$, $n = 198$, and $n = 99$ for EtOH stored samples, respectively, and $n = 557$, $n = 371$, $n = 278$, $n = 223$, and $n = 111$ for frozen stored samples). Calibration samples were selected by sample number: every other sample was used as calibration for 0.5, every third sample was used for 0.33, every fourth sample used for 0.25, every fifth sample used for 0.2, and every tenth sample used for 0.1. As a result, the specific samples utilized for lower frequency calibration sets did not necessary belong to higher frequency calibration sets (e.g., the fifth sample was used for calibration of the 0.2 frequency test, but was used as a test sample for the 0.5 frequency test). Unlike the leave-one-out analysis explained above, here a PLSR model was created using only those samples marked as belonging to the calibration set. All test set samples were then plotted along this regression, with goodness of fit being recorded in terms of R^2, RMSEP (root mean square error of prediction), and RPD. These values were compared across each of the calibration set frequencies.

As previous studies have found distinctive differences in spectra produced from intra-species samples from different environments (Wedding et al. 2014; Robins et al. 2015; Helser et al. 2019), it was hypothesized that there would be a non-linear relationship in both spectra-age correlations and spectra-length correlations between the day 1-20 samples and the day 24-49 samples, corresponding to separate collection environments. Following collection, the discontinuous age-length relationship between these sets reinforced this possibility. Accordingly, these samples were split into 2 distinct groups by storage set (day 1-20, n = 425; day 24-49, n = 566 for EtOH stored; day 1-20, n = 577; day 24-49, n = 538 for frozen stored). Each group was spectrally optimized and modeled independently in a leave-one-out analysis, as well as in the 0.5, 0.33, 0.25, 0.2, and 0.1 frequency calibration set test validations. For the day 1-20 set, a spectral optimization method utilizing a first derivative transformation (with 17 smoothing points) and vector normalization between wavenumbers 9400 – 6096 cm^{-1} and wavenumbers 5456 – 4248 cm^{-1} was used in the EtOH stored group; in the frozen stored group, vector normalization in the 9400 – 5448 cm^{-1} wavenumber range was found to be best. For the day 24-49 set, a first derivative transformation (with 17 smoothing points) in the 9400 – 7496 cm^{-1} and 6104 – 5448 cm^{-1} wavenumber ranges was used in the EtOH group; a 1^{st} derivative transformation and vector normalization in the 9400 – 6096 cm^{-1} and 5456 – 4248 cm^{-1} wavenumber ranges were used on the frozen sample set.

After reviewing the results of the above models, it was determined that the spectra produced by samples 1-10 days old were aberrant within all PLSR models created which included them. Accordingly, a leave-one-out analysis was performed using only samples from days 12-49 (n = 789 for EtOH; n = 754 for frozen). The spectra were optimized as

described above, and a new model was created. Tests of introduced error, described below, used the EtOH stored age 12-49 set to calculate and compare to.

In order to estimate the amount of error which would have been introduced by utilizing traditional age estimations for our calibration and test samples, two methods were tested. First, the light microscopy results of Jones and Brothers' (1987) study examining daily otolith-increment deposition in *M. saxatilis* were used. As the error associated with otolith-based ageing was dependent on food availability (which was not precisely known for these samples, and foraging success might have varied by individual), I constructed 5 hypothetical input ages for each of my samples. These corresponded to the 5 feeding conditions examined by Jones and Brothers (1987): Always fed, Starved, Starved/fed, Intermittent, and Always fed (using only fish ≤68 days old). The input age was calculated for each of these using the equations presented by Jones and Brothers (table 2.2). Assuming a normal distribution for both slope and intercept, the mean and regression coefficient of each feeding group were used to generate a unique age prediction equation for each sample (EtOH stored, ages 12-49), for each feeding treatment, using R (version 3.5.2). The equations for each treatment were then used to generate simulated erroneous age predictions using the true age of each sample. A standard value of 4 was added to each prediction, as it was determined that in normal conditions initial increment formation began at 4 days (Jones and Brothers 1987). These predictions were then rounded to the nearest whole number to best approximate traditionally estimated ages, with linear regression models created to ensure the R^2 value was still equivalent to the original.

Table 2.2. Effects on daily otolith-increment deposition of different feeding regimes (reproduced from Jones and Brothers 1987).

Condition	Slope (increment counts/ day)	Intercept	R^2
1. Always fed	0.946	-3.627	0.96
2. Starved	0.469	-1.697	0.77
3. Starved/ fed	0.930	-10.430	0.90
4. Intermittent	0.873	2.579	0.96
5. Always fed (≤68 days)	0.980	-4.016	0.96

These generated age predictions, simulating estimated ages under 5 hypothetical feeding regimes, were used to create new FT-NIRS PLSR models. Spectral data for each sample was copied, with the associated age changed to match that of the simulated prediction age for each condition. These new sets were spectrally optimized, as described above, and used to generate a leave-one-out analysis for each of the 5 conditions (table 2.2).

To allow for further comparison with traditional studies of ageing accuracy, modifications were made to the average percent error (APE) calculation of Beamish and Fournier (1981). While the original equation includes measures of precision between multiple readings of an ageing structure, this does not apply to FT-NIRS ageing, where only 1 age estimate is made per model created. Additionally, using known age fish eliminates the need to use average age: known age is used instead. The modified equation is defined as follows:

N is the total number of fish aged, X_{Ij} is the predicted age for fish j, and X_j is the known age of fish j.

$$\frac{1}{N}\left[\sum_{i=1}^{N}\frac{|(X_{ij}-X_i)|}{X_i}\right]$$

As the error found in traditional ageing studies is often reported as APE, our second measure of the error introduced by using traditionally aged fish was to create hypothetical age models using simulated ages generated by varying APE. This was done for all EtOH stored day 12-49 samples. For each percentage of APE examined (ranging from 1-10), random normal distributions were created. A standard deviation of 1/3rd of APE was assumed, and the resultant "error" to apply to each known age was randomly assigned as either positive or negative. Values were then rounded to the nearest whole number, to simulate predicted ages. Each APE was rechecked following rounding, and age estimates which now differed more than 0.1 from the approximated APE were reperformed (note: for the APE of 2, the standard deviation had to be increased to 1.66 to result in a post-rounding APE of 2.0). For 1% APE, rounding of the half ages (those aged 24-49) provided the majority of the 1% average error in the set. Ages generated for each simulated APE were input as known ages, and new FT-NIRS PLSR models were optimized and created for each, again using a leave-one-out cross validation. RMSECV, %RMSE, and APE (using true ages as well as model-predicted ages) were calculated for each. This entire process was repeated, except each treatment was assigned positive error (systemic age overestimation) or negative error (systemic age underestimation) rather than randomly assigning positive or negative error estimations per sample.

The standard lengths for each sample were plotted against known age, at first without inclusion of spectral data. A regression line was created using this data, with the strength of correlation measured in R^2. Use of a von Bertalanffy function was rejected, as samples did not approach asymptotic (or even strongly non-linear) growth within this age range. The equation generated for this regression was used to calculate a per-sample error in prediction. This in turn was used to calculate a root mean square error of prediction (RMSEP), for ease of comparison with spectral models. Length data was then combined with spectral data to produce a PLSR model, as was utilized for age. The most appropriate spectral preprocessing was again tested, this time for the spectra-length relationship. In this case a pretreatment utilizing a first derivative transformation (17 smoothing points) between 6104 – 5448 and 4600 – 4248 cm^{-1} was found to be the most effective. A leave-one-out analysis was performed, as described previously.

Deviations were then compared, to test whether days with high variation in fish length corresponded to the days with high variation in test prediction accuracy. Variation and standard deviation were both compared to test whether a correlation existed between these. Specific outliers for each group were compared, to test whether any were shared between spectra-length, spectra-age, or age-length models.

Results

Test set validations using 0.5, 0.33, and 0.25 frequencies were found to be very comparable in the all-ages group across both R^2 and RMSEP, with minor drops in accuracy seen in the 0.2 and 0.1 calibration frequency sets (RMSEP increases of 0.12 and 0.13 in the EtOH set and 0.31 and 0.51 in the frozen set, respectively). The combined

ages EtOH stored leave-one-out spectra-age model was also more accurate at predicting age than was the length-age model: the R^2 and RMSECV values for the leave-one-out analysis were 92.79 and 3.97, while the R^2 and RMSEP values for the length-age correlation were 89.8 and 4.99. In other words, the leave-one-out spectra-age model had a prediction accuracy higher by approximately 1 day than that of the length-age regression. The differences in predicted age for each sample, and the differences in error of prediction per sample, were both significant (χ^2, $p < 0.05$). All calibration set frequencies for the combined-ages models produced more accurate age predictions than was found for the length-age model (table A.1).

EtOH stored

The day 1-20 set was less accurate than those of the combined set and the day 24-49 set, in the leave-one-out analyses and all test set validations (figures 2.1, 2.2, and 2.3). The leave-one-out analysis for the ages 1-20 group was slightly more accurate (R^2 = 82.65 and $\%RMSE_{Range}$ = 12.70) than that for the ages 24-49 group (R^2 = 80.46 and $\%RMSE_{Range}$ = 13.12) in terms of the range of dates, though the %RMSE values of the day 20-49 set were superior to the others, corresponding to the increased maximum age. The APE difference between the two was large (6.82% for days 24-49 vs. 43.7% for days 1-20) and in favor of the day 24-49 set. Test sets for the ages 24-49 group were all more accurate than those of the ages 1-20 group in terms of %RMSE and APE as well (table A.1). Spectral loadings for both groups contained significantly more “noise” than was seen in that of the combined set (figures 2.7, 2.8, and 2.9).

In both the all-ages and day 1-20 sets, the leave-one-out cross validation model was found to be the most accurate. For the day 24-49 set, the test set model utilizing a sample frequency of 0.5 as a calibration set was more accurate than the leave-one-out analysis in R^2 and RMSE. A general trend can be found in accuracy across sets, with higher calibration set frequencies producing more accurate models. Exceptions include the 0.5 calibration set frequency model for combined ages, as well as the 0.5 and 0.33 calibration set frequency models for the days 24-49 group.

Removing days 1-10 from the leave-one-out cross validation (using days 12-49 only, $n = 789$) resulted in a significantly more accurate model (figure 2.5). The RMSECV was decreased by 0.67 days in the reduced set (a change from %RMSE of 8.02 to 6.69), and the APE was significantly improved (33.04% to 9.2%). By APE, this new set was second in accuracy only to the day 24-49 group.

For all but 2 of the combined set age-spectra models, a rank 7 regression was found to be the best at maximizing accuracy and minimizing the inclusion of spectral noise. A rank 7 regression was likewise found to be most appropriate for 2 of 6 of the day 1-20 models, and for 5 of 6 of the day 24-49 models (table A.1). A rank 8 regression was utilized in the age 12-49 leave-one-out validation. The majority of the APE simulations likewise used a rank 8 regression.

Ages which displayed a high variation in length were more likely to display a high variation in the error of spectral age prediction. No significant difference was found between average standard deviations, per day, of length or spectra model generated error of prediction (χ^2, $p>0.05$). However, significant difference was seen when examining

only the first 20 days (χ^2, $p<0.05$) (figure 2.6). The similarity in deviation between day 24-49 length and prediction errors can be seen in figure 2.4, as can the dissimilarity of the same between day 1-20.

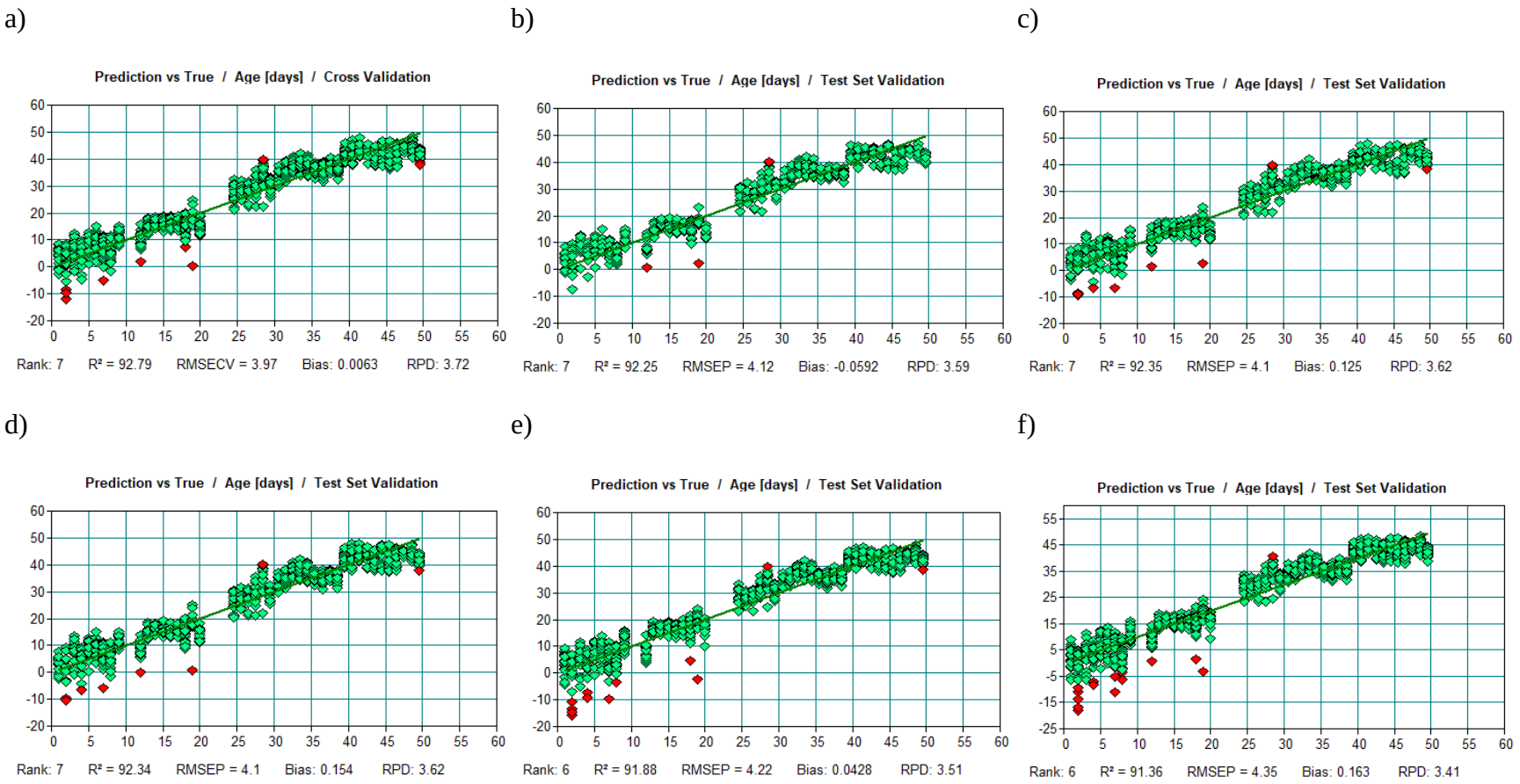


Figure 2.1. All ages, a) leave-one-out cross validation, b) 50% calibration set test validation, c) 33% calibration set test validation, d) 25% calibration set test validation, e) 20% calibration set test validation, f) 10% calibration set test validation.

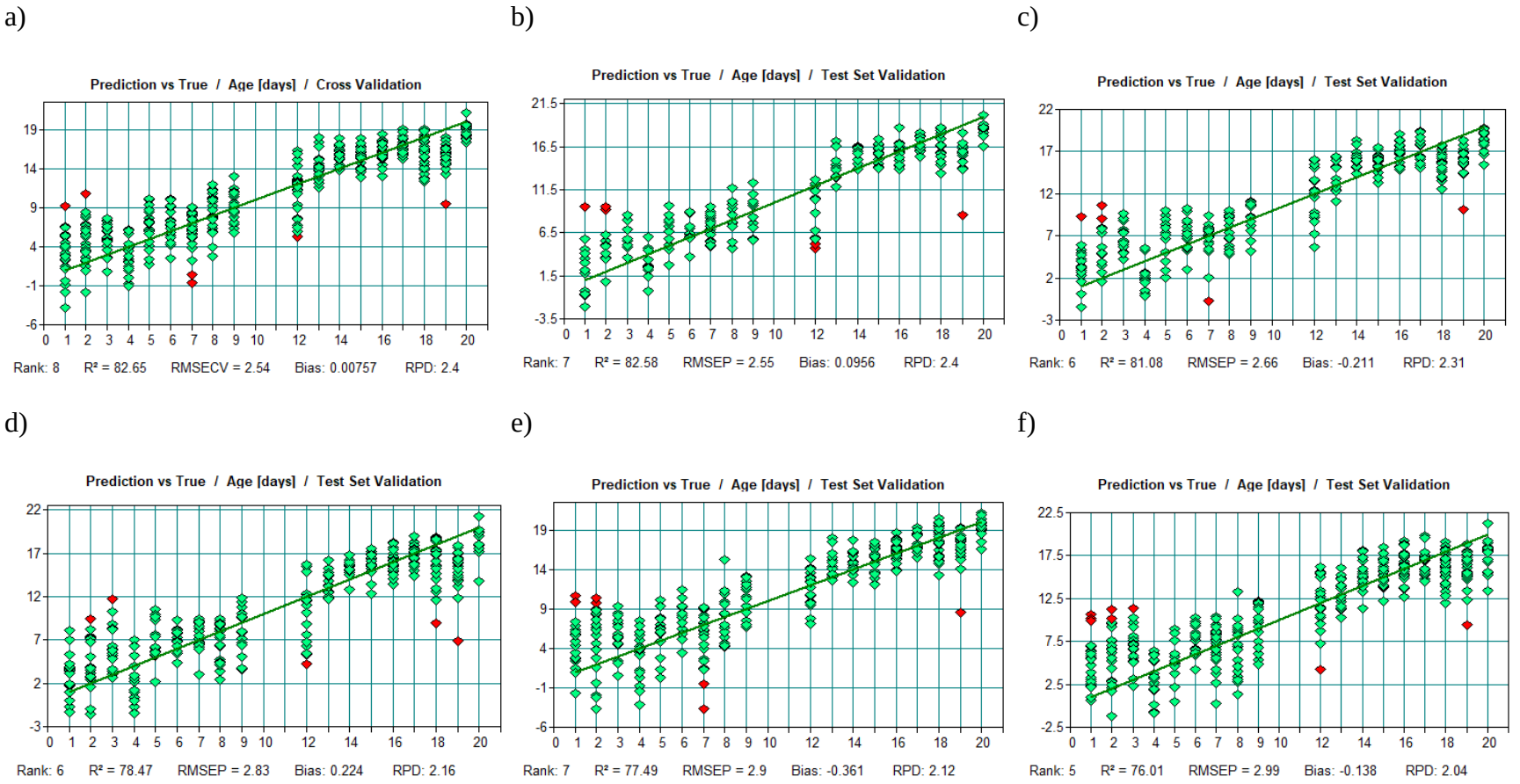


Figure 2.2. Days 1-20, a) leave-one-out cross validation, b) 50% calibration set test validation, c) 33% calibration set test validation, d) 25% calibration set test validation, e) 20% calibration set test validation, f) 10% calibration set test validation.

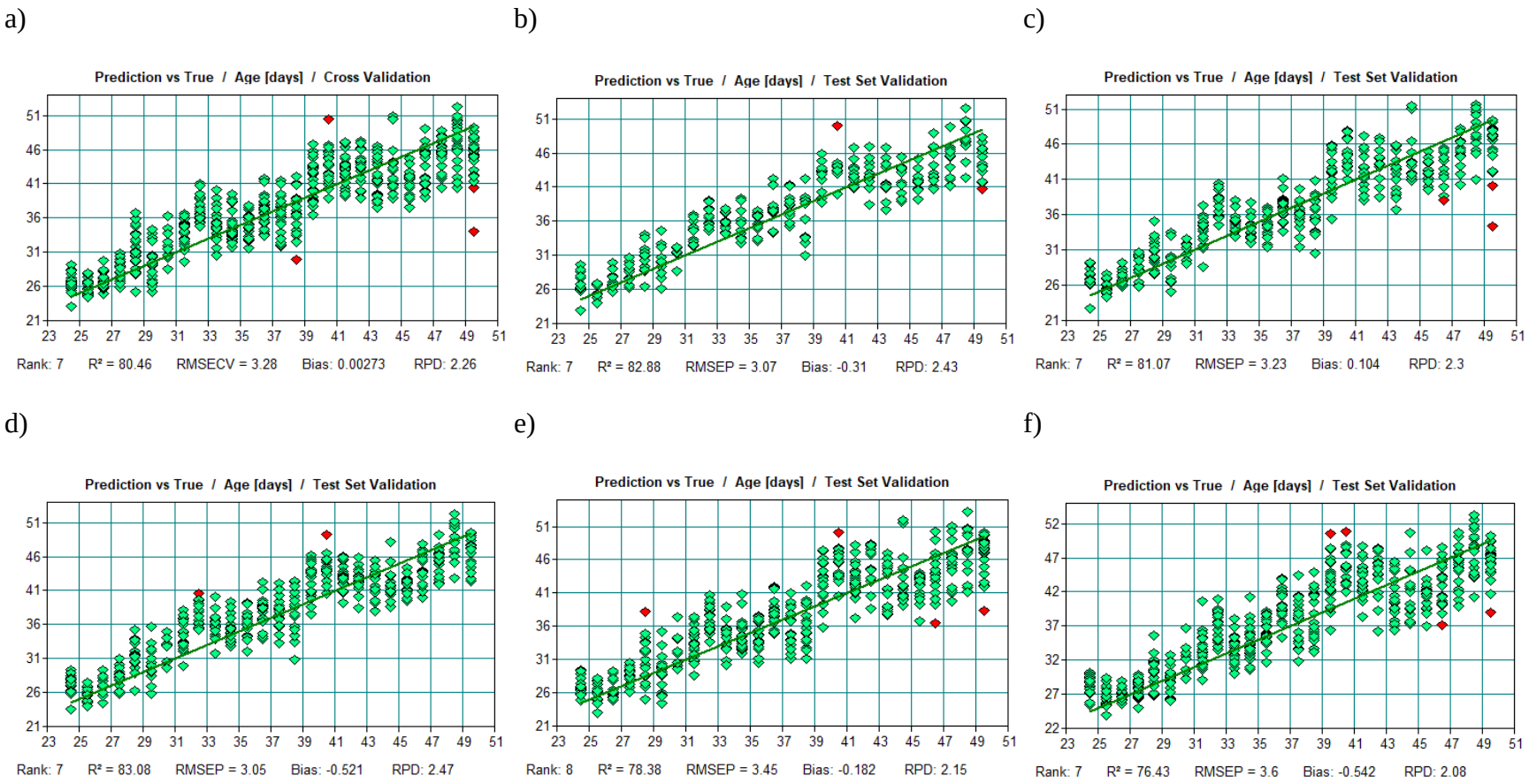


Figure 2.3. Days 24-49, a) leave-one-out cross validation, b) 50% calibration set test validation, c) 33% calibration set test validation, d) 25% calibration set test validation, e) 20% calibration set test validation, f) 10% calibration set test validation

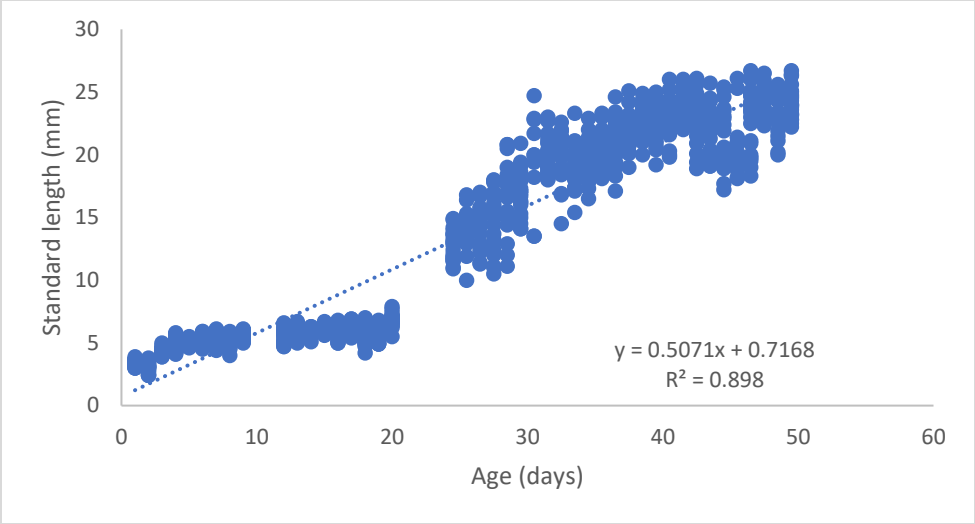


Figure 2.4. Lengths plotted by age, with regression line and equation. The equation of this relationship was used to calculate the length-predicted ages for each sample, the RMSE of which is displayed in table A.1.

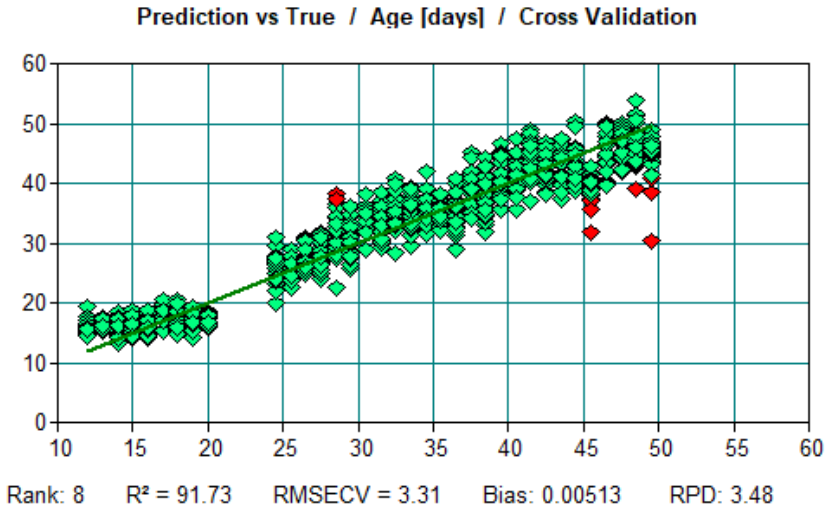


Figure 2.5. Days 12-49 leave-one-out cross validation results.

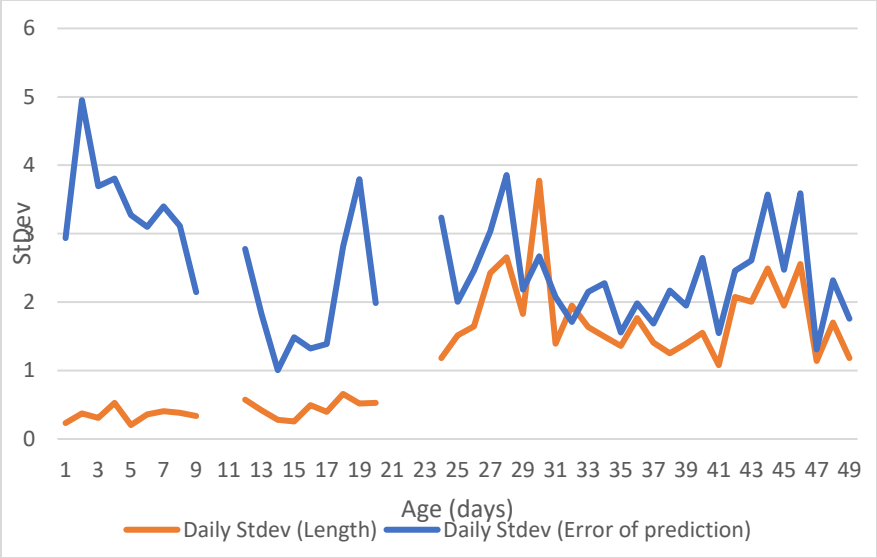


Figure 2.6. A comparison of standard deviations for length and spectral error of prediction, by age.

The spectra-length model was very accurate at predicting sample length, with an R^2 of 98.32 and a RMSECV of 1.03 (3.86% of maximum sample length) (figure 2.11). Prediction accuracy was generally greater with smaller lengths, with the greatest density of deviation from predicted lengths seen in samples with standard lengths between 18 and 26 mm. The gap of samples between 7.9 mm and 10 mm corresponds to the gap between the age 1-20 group and the age 24-49 group. This trend corresponds with the trend for days past the initial 20 to have higher deviations in length (figure 2.4).

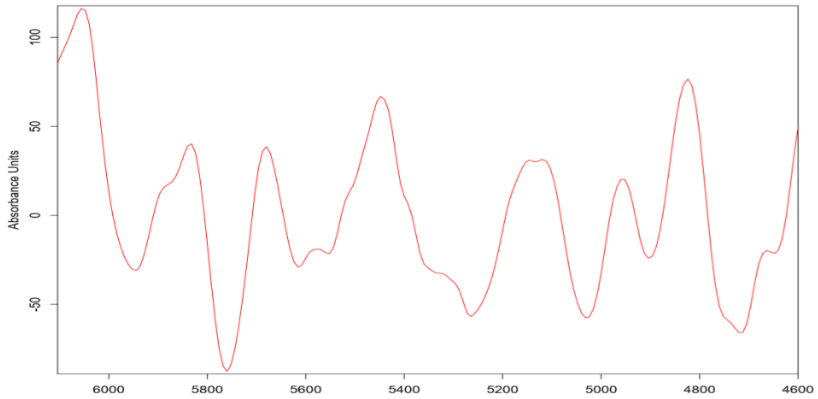


Figure 2.7. Spectral regression coefficient loading for combined all-ages set.

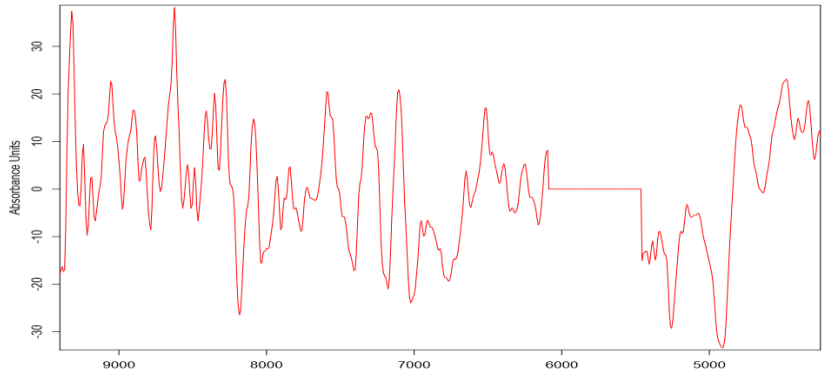


Figure 2.8. Spectral regression coefficient loading for ages 1-20 set.

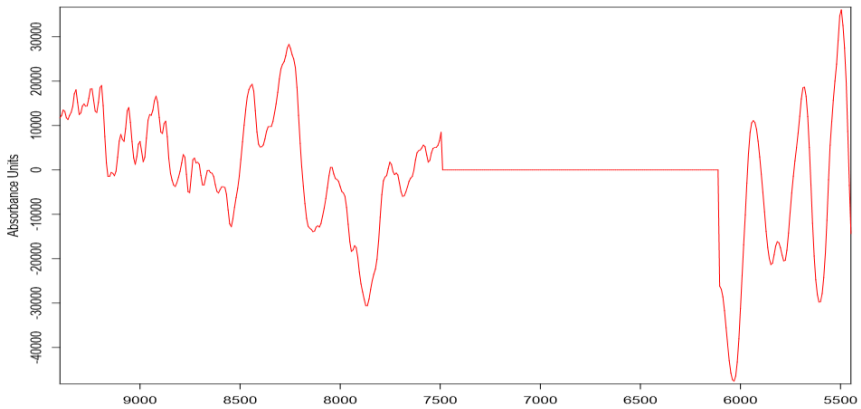


Figure 2.9. Spectral regression coefficient loading for ages 24-49 set.

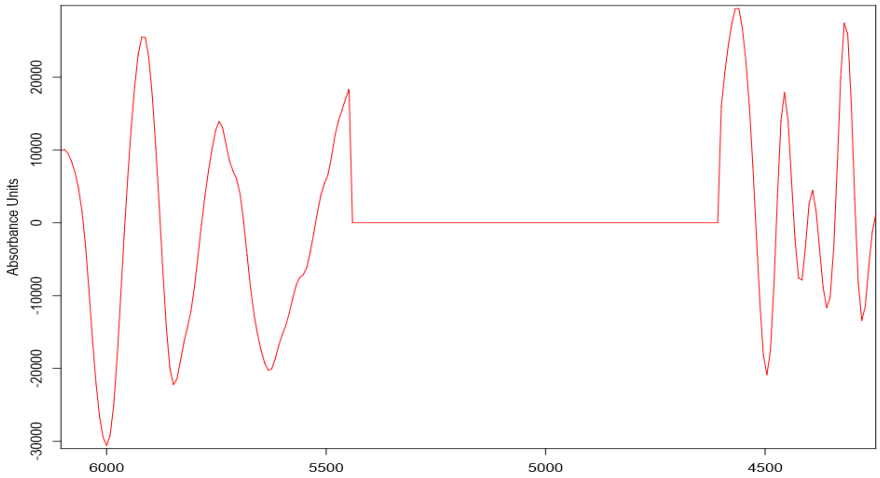


Figure 2.10. Spectral regression coefficient loading for the spectra-length model.

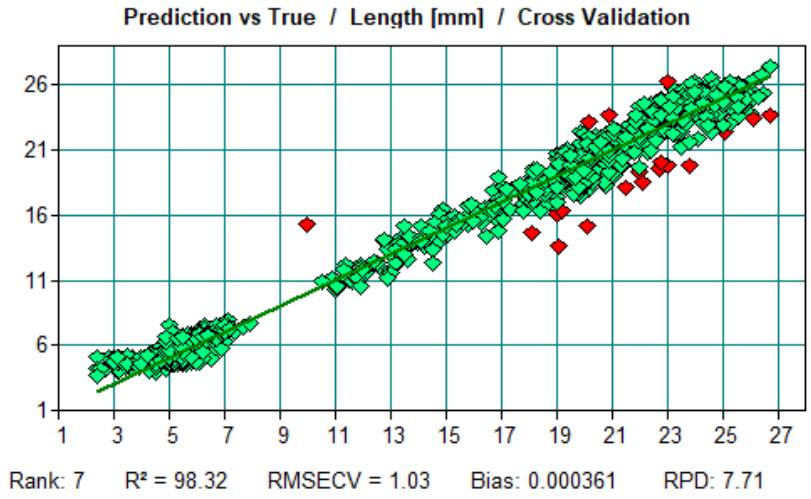


Figure 2.11. Spectra-length cross validation model, utilizing a leave-one-out analysis.

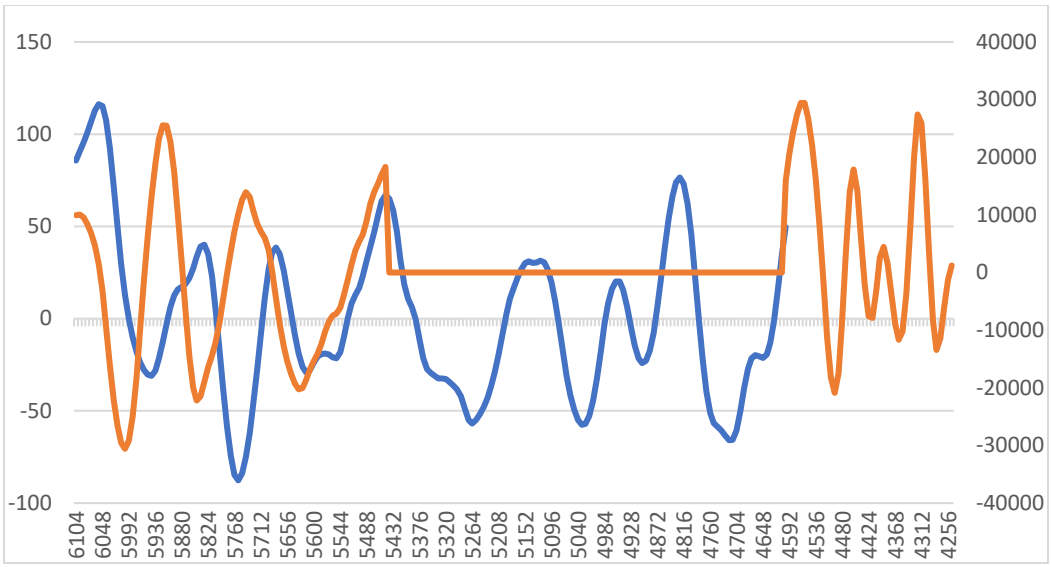


Figure 2.12. Spectral regression coefficient loadings for the spectra-length model (orange) and the spectra-age model (blue)

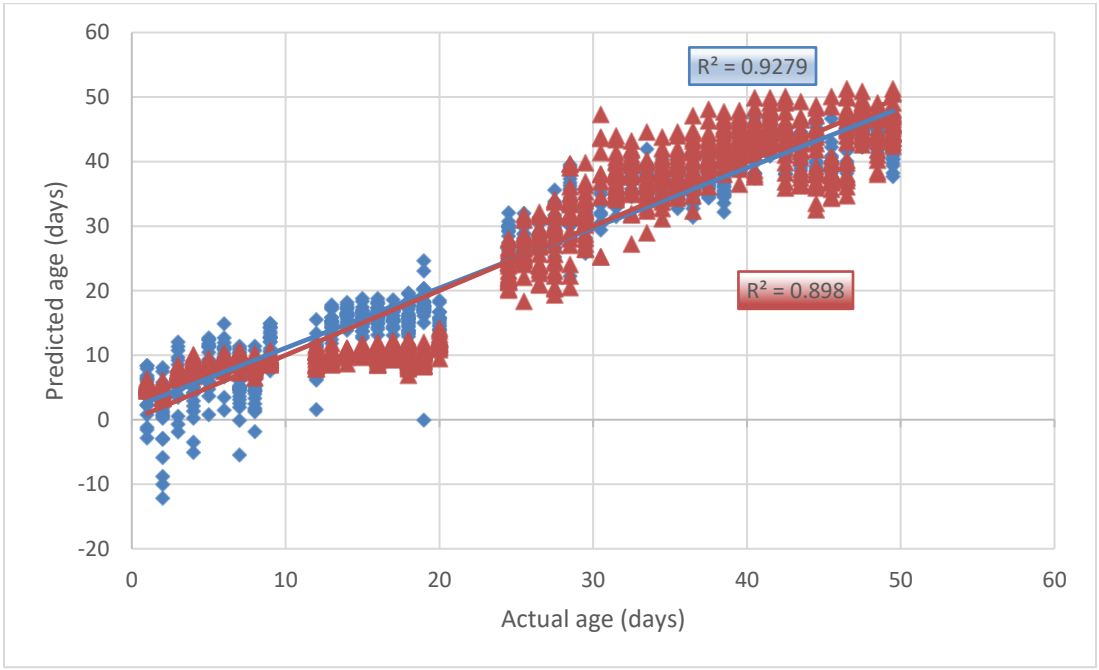


Figure 2.13. Predicted ages for the length-age model (red) and the spectra-age model (blue) vs. actual age with linear regression lines for each model.

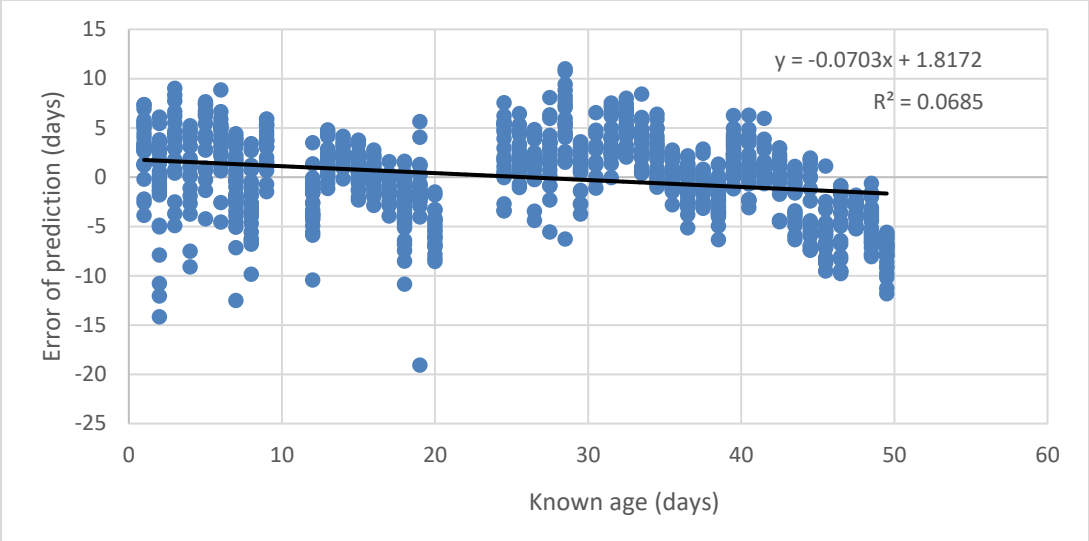


Figure 2.14. Cross validation model error of prediction (in day) vs. known ages.

Table 2.3. Predictions aged to within a given number of days for each sample in a leave-one-out cross validation model.

Set	%within 2 days	%within 3 days	%within 5 days
All ages	49.55	62.87	83.25
Ages 12-49	55.26	71.74	91.51
Ages 1-20	19.76	29.18	43.76
Ages 24-49	57.95	71.38	90.64

Table 2.4. The percentage of samples which were accurately predicted to within either 10, 15, 20, 25, or 30% of true age.

Set	% accurate to 10%	% accurate to 15%	% accurate to 20%	% accurate to 25%	% accurate to 30%
All ages	47.73	62.97	72.05	78.30	81.03
Ages 12-49	62.73	82.26	90.75	95.31	97.59
Ages 1-20	38.59	50.59	60.24	66.59	71.06
Ages 24-49	75.87	92.22	97.70	99.29	99.82

Each leave-one-out analysis group (combined, ages 12-49, ages 1-20, and ages 24-49) were examined for the percentage of samples predicted to within 2, 3, and 5 days

of true (table 2.3). Only the sets excluding day 1-10 fish had accurate predictions to within 2 days in more than half of samples. These sets also were able to predict ~70% of samples to within 3 days, and ~90% to within 5 days. While slightly less accurate, the combined set was likewise able to place most samples to within 3 days. The group exclusively utilizing young fish (ages 1-20) was unable to place half of the samples within 5 days of true age.

Likewise, the sets excluding days 1-10 had >62% of samples accurately aged to within 10% of true age, >82% within 15% of true, and >90% within 20% of true. The combined ages set and the ages 1-20 set were able to accurately predict most samples to within 15% of true, though the ages 1-20 set was only able to place ~70% of samples to within 30% of true age (table 2.4). The ages 24-49 set had the most accurate age predictions across samples, which aligns with the average APE comparison (table A.1).

Leave-one-out models created using the results of Jones and Brothers (1987) feeding experiment showed highly variable predictive accuracy. The always fed treatments (1 and 5) performed best in terms of APE from true (9.22 and 8.79, respectively). The starved, starved/ fed, and intermittently fed models were all significantly more inaccurate than any other in APE from true, but the intermittent treatment had relatively low APE from the input simulated ages.

The strength of models utilizing introduced error (in the form of APE) largely depended on error directionality (table A.2). Models which utilized a randomly assigned positive or negative error for each sample were more likely to perform poorly in measures of model accuracy (including R^2 and RMSECV). However, they were also

more likely to maintain a relatively low APE to true ages and a higher APE to input ages. Conversely, models created utilizing consistently underaged or overaged fish were likely to appear more successful in measures of model accuracy (R^2 and RMSECV) but had higher APE to true and lower APE to simulated (figure 2.15). This trend became more pronounced as the amount of introduced error increased.

Table 2.5. Leave-one-out model results using treatment 1-5 parameter generated age sets (from Jones and Brothers, 1987).

Treatment	n	R2	RMSECV	APE(true)	APE(sim)	Rank	RPD
Treatment 1- Always fed	789	88	3.81	9.22	11.04	8	2.89
Treatment 2- Starved	789	71.21	3.3	43.49	15.88	7	1.86
Treatment 3- Starved/ fed	789	82.8	4.68	30.97	26.27	7	2.41
Treatment 4- Intermittent	789	88.34	3.5	16.05	8.39	8	2.93
Treatment 5- Always fed (young only)	789	88.82	3.85	8.79	10.65	8	2.99

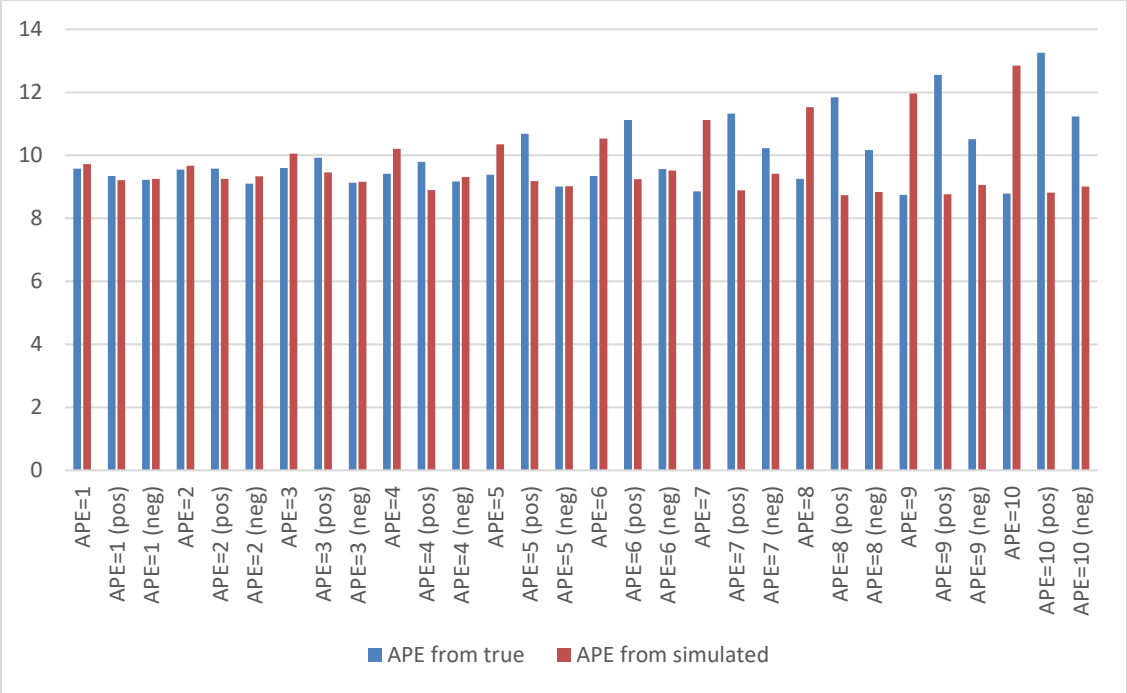


Figure 2.15. APE of model-generated ages from true ages and simulated ages for artificially generated 1-10% APE input ages. Ages were randomly assigned as positive or negative in the main set; (pos) indicates that the artificial error was positive (over ageing), while (neg) indicates that artificial error was negative (under ageing).

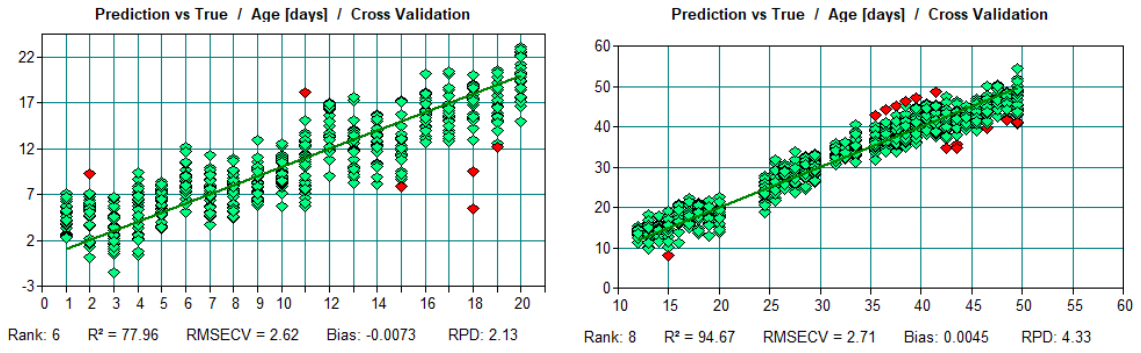
Frozen Stored

Similar to what was seen with the EtOH stored samples, accuracy in the frozen samples was generally higher with increasing calibration set frequencies in terms of both R^2, RMSEP, and APE. Exceptions to this trend were found in the all ages 0.33 calibration set model (with a higher R^2 and lower RMSEP than the 0.5 calibration frequency model), in the all ages 0.2 calibration set model (with the 0.2 model having a lower RMSEP and APE than the 0.25 calibration set) and in the age 24-49 0.2 calibration set model (which had a higher R^2, lower RMSEP, and lower APE than the 0.25 calibration set frequency model). Unlike the EtOH stored set models, the 0.5 calibration frequency models were generally more accurate than the leave-one-out validation analyses.

The age 12-49 group leave-one-out analysis, excluding the youngest individuals, was the most accurate model produced, with an R^2 of 94.67, a RMSE of 2.71, and a %RMSE of 5.53. The RPD produced for this model was the highest of those produced using frozen samples at 4.33. The least accurate models were produced using the age 1-20 set, with each having an $R^2 \leq 80.33$ and a %RMSE > 12. While the combined age group did have some signs indicative of strong FT-NIRS models ($R^2 > 90$, %RMSE ~ 8, RPD > 3), the APE (>32%) and % within tests (table A.3) showed it to be highly inaccurate.

Frozen stored sample models were generally slightly more accurate than those produced by samples stored in EtOH by the metrics of model success (R^2 and RMSE), but not consistently by APE. The combined age set in particular had a lower APE in the EtOH samples when compared to the frozen samples. The younger age set had lower RMSE and APE in the frozen samples, as did the older age set. Differences between models of comparable age and calibration size sets between storage media were generally minimal (R^2 difference usually < 2, RMSE differences < 0.6).

a) b)



c) d)

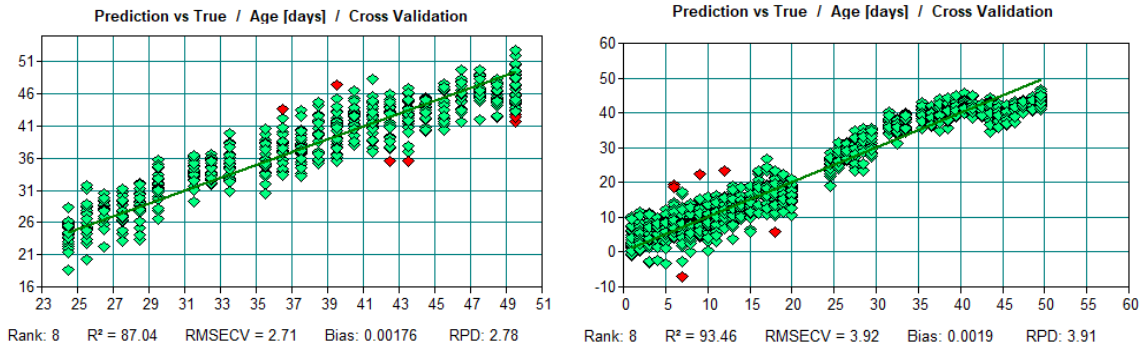


Figure 2.16. Regressions created from leave-one-out analyses using frozen samples of a) ages 1-20, b) ages 12-49, c) ages 24-49, and d) all ages.

Table 2.6. Percent within *x* day results for each frozen stored set.

Set	%within 2 days	%within 3 days	%within 5 days
All ages	44.30	60.27	84.57
Ages 12-49	65.11	82.10	95.76
Ages 1-20	66.78	84.43	96.71
Ages 24-49	64.13	81.04	94.98

Combined Storage

a) b)

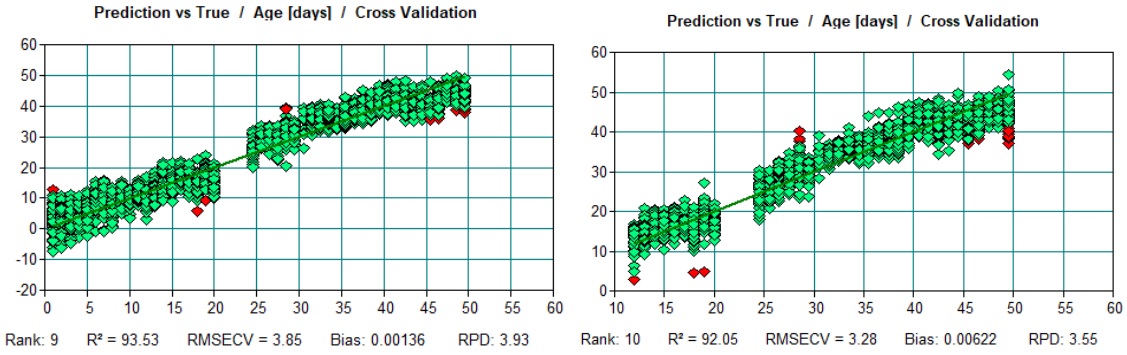


Figure 2.17. Leave-one-out cross validation results of combined storage sets, including a) all ages, and b) ages 12-49.

The two leave-one-out validation models created using both EtOH and frozen stored samples were largely accurate. The set including all ages had a RMSECV less than either the EtOH or frozen stored sets had separately, and resulted in age predictions which were marginally more accurate as well (APE = 32.65 vs. APE = 32.9 and 33.04). The APE produced from the ages 12-49 combined storage set was slightly higher than those produced from the EtOH and frozen sets (APE = 9.4 vs. 9.2 and 7.65, respectively). Ranks used for these combined storage models were higher than those used for any other model created

Table 2.7. Leave-one-out cross validation analyses using combined storage sets.

Test	*n*	R^2	RMSECV	%RMSE	APE	Rank	Bias	RPD
All ages	2106	93.53	3.85	7.8571	32.65	9	0.0013	3.93
Ages 12-49	1543	92.05	3.28	6.6938	9.4	10	0.0062	3.55

Similar to the separate storage models, the leave-one-out validation analysis using all ages was unable to place half of the samples to within 2 days of true age (table 2.7). The model created using only ages 12-49 was able to place most samples to within 2 days, and 91% to within 5 days (similar to that produced using EtOH stored samples only and slightly less than that produced using only frozen samples).

Table 2.8. Percent within *x* day results for combined storage sets

Set	%within 2 days	%within 3 days	%within 5 days
All ages	46.68	61.92	84.24
Ages 12-49	57.29	73.04	91.32

Discussion

That the combined spectra -- age model was more accurate a predictor of age than the length – age model is a strong defense of continued exploration into the utility of NIRS as a fisheries ageing tool. As the tested sample set contained excellent candidates for length-based ageing, being young and fast growing (Campana 2001), this comparison was likely one of the most difficult NIRS ageing protocols could have encountered with regard to relative accuracy of length-based prediction. As the first model created utilizing NIRS for bony fish ageing without any uncertainty in calibration or test age input, these results are the first for which no results can be assigned to traditional ageing method errors.

These findings are also the first which suggest the usage of whole-fish might be a viable strategy for future studies. As is seen in the split age models, the reduction in age range influences prediction accuracy in a way which is not wholly counteracted by

dividing prediction error by the age range of a sample set. The frozen stored age 12-49 group %RMSE of 5.67 is less accurate than that found for *Lutjanus campechanus* otoliths (%RMSE = 5.06, Passerotti et al. 2020a) and *Gadus chalcogrammus* otoliths (%RMSE = 4.87, Helser et al. 2019). However, it showed improved percent accuracy over *Squalus megalops* dorsal fin spines and fin clips (%RMSE = 9.64 and 10.7, Rigby et al. 2014), *Squalus megalops* vertebrae (%RMSE = 7.40, Rigby et al. 2014), *Squalus montalbani* dorsal fin spines (%RMSE = 9.54, Rigby et al. 2014), *Sphyrna mokarran* vertebrae (%RMSE = 8.52, Rigby et al. 2015), *Lutjanus malabaricus* otoliths (%RMSE = 5.87, Wedding et al. 2014), *Lates calcarifer* otoliths (%RMSE = 6.25, Robins et al. 2015), *Pagrus auratus* otoliths (%RMSE = 6.12, Robins et al. 2015), and *Carcharhinus sorrah* vertebrae (%RMSE = 8.97, Rigby et al. 2015). Looking at predictive accuracy as an absolute, the error found here is the lowest yet published in any NIRS fish ageing study, with a combined set RMSE of less than 4 days and an older subset (12-49) RMSE lower than 3 days. This suggests that one limit of NIRS prediction accuracy to date might be the limitation on developing a robust fine-resolution calibration set imposed by the reliance upon traditionally ageing the calibration samples, as well as having less well distributed sample ages.

Previous studies have found different wavelength regions to be the most important for age correlation, depending on taxa and the structure investigated. Helser et al. (2019) found the 6821 – 5269 cm^{-1} and 5022 – 4171 cm^{-1} regions to be most relevant in walleye pollock (*Gadus chalcogrammus*) otoliths. Rigby et al. (2014) found the areas between 9300 – 8200 cm^{-1}, 7800 – 6800 cm^{-1}, and 4600 – 4000 cm^{-1} to be the most strongly correlated to age across all species and structures examined (*S. megalops*, *S.*

montalbani, *A. pallidus*; vertebrae, fin clips, and dorsal fin spines). Robins et al. (2015) utilized the 4832 – 4327 cm^{-1} region for *Lates calcarifer* otoliths and the 6160 – 4580 cm^{-1} region for *Pagrus auratus* otoliths. Passerotti et al. (2020b) utilized the 7506 – 4242 cm^{-1} region in ageing juvenile *Lutjanus campechanus* otoliths. In the present study, it was found that the 6104 – 4600 cm^{-1} region was most useful for the combined ages model of whole fish stored in EtOH, and the 9200 – 4248 cm^{-1} region was most important for frozen stored fish. The discontinuous nature of many of these spectral regions suggests that any single molecular bond is unlikely to be responsible for the FT-NIRS – age relationship across taxa and structure. Even using the same structure researchers have found different key spectral regions between taxa (Helser et al. 2019; Robins et al. 2015; Wedding et al. 2014; Passerotti et al. 2020a). Wedding et al. (2014) found that the key spectral frequencies varied even within the same species and attributed this to differences in season or geographical range. While the identification of key molecules utilized by FT-NIRS to correlate with age might help in planning future research efforts, studies into the ability of FT-NIRS to predict age will likely need to be performed for each potential fishery to be analyzed regardless.

The conclusion that numerous factors can influence which wavenumbers are the most relevant to age-spectra relationships is reinforced from results found here. Whether samples were stored in EtOH or frozen changed the wavenumbers which were utilized in model creation. This is consistent with the theoretical understanding of NIR spectra for several reasons. First, the molecular environment around a sample can cause a shift in spectral peaks produced by changing the vibrational frequencies of a given bond (Miller 2001). Temperature can also cause a shift in spectra produced by a given molecule, which

would result in frozen samples producing distinctive peaks for the same bond (Miller 2001). This is illustrated in figure 2.18 below, which displays the spectra produced from a single bond in water and ice at varying temperatures. Frozen samples which did not reach the same ambient temperature as those stored in EtOH would therefore be expected to produce slightly different spectra even without any difference in the composition of storage media. Tests using a small number of frozen samples at various stages of thawing found only small shifts in spectra produced, but as these shifts are not equally distributed throughout the spectrum direct similarities between frozen and non-frozen samples are difficult to directly attribute to this.

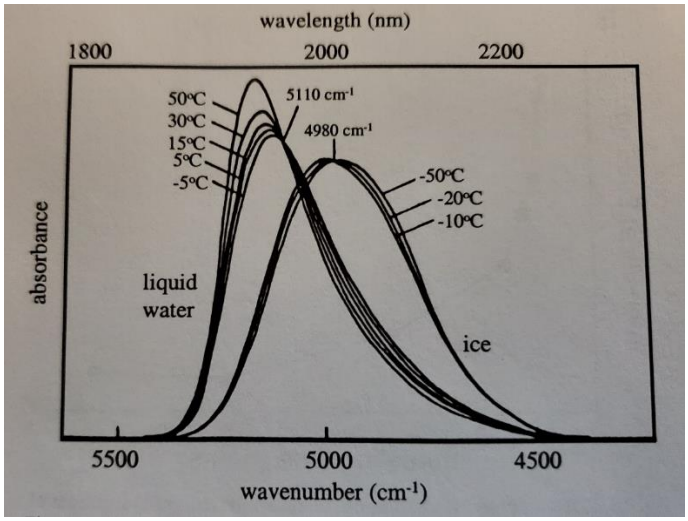


Figure 2.18. Near-infrared spectra produced from water and ice at varying temperatures (from Fornes and Chaussidon 1978).

The fact that some models using fewer calibration samples were more accurate than those using more has important implications for FT-NIRS ageing studies moving forward. After investigation, it was found that the primary reason for this was the inclusion of spectral outliers in some calibration sets, while they were included in the test sets for other models. For instance, in the EtOH stored age 24-49 set, the model using a

calibration frequency of 0.25 of all samples predicted age marginally more accurately than the model using 0.5 of all samples for calibration (RMSE = 3.05 vs. 3.07). As the calibration samples were chosen randomly within each age group (by virtue of arbitrary sample numbering), models were not as optimized as they could have been. After manually including/ excluding all outliers in the calibration set in turn, a change in R^2 was seen between -0.5 and 0.43, and a change in RMSEP was seen between -0.27 and 0.43. While the random selection of calibration samples is appropriate for examining the potential of FT-NIRS in a fisheries application (where only samples intending to be used for calibration will have traditional ages established), it does suggest that comparisons between FT-NIRS studies will need to consider the way in which samples were assigned to calibration or test sets. Furthermore, it limits interpretation of small differences in model success.

A critical threshold of calibration set size does seem to be present in all three age groups for which EtOH stored test models were created (combined, 1-20, 24-49). The most significant drop in accuracy in all groups occurs when the calibration set size is reduced to below 140 samples. Interestingly, this trend is weakest in the combined age set, despite it having the lowest calibration set size at this threshold (n = 99), spread over the greatest range of ages. This is perhaps due to the greater biological variability still present within the combined age set relative to the split age groups, which reinforces the conclusion expounded previously that greater sample variability might have higher importance than calibration set size. Despite only having approximately 2 calibration samples per age group in the 0.1 calibration set frequency model, the combined ages set showed greater relative predictive accuracy (%RMSE = 8.79) than the model created

using approximately 10 samples per day in the younger split age group (%RMSE = 12.75).

I found significantly worse predictive capabilities for fish aged 1-10 days across all models which included them. While care must be taken to ensure that aberrant age are not simply the result of a regression to the mean (the importance of which in FT-NIRS ageing is discussed in Passerotti et al. 2020b), there exist a number of strong, mathematically independent reasons to suggest this group might not be well predicted by the entire set. First, the presence of the yolk sac, visible to the FT-NIRS aperture by my positioning, was only found in fish aged under 5 days old (though some trace could still be seen in those as old as 9 days old). This molecularly unique feature in our samples would be expected to frustrate spectra-age correlations by presenting absorbance frequencies which are not present to any degree in later ages. Second, the beginning of exogenous feeding around day 5 would be expected to change the molecular composition of each sample's skin, causing further spectral discontinuity between these ages. Third, the small physical size of these samples (often <4 mm L_{St}) meant that sample positioning varied more in this group than any other. As sample positioning changed the structures facing the aperture, inconsistency brought about by difficulty manipulating these small samples would likely confound any molecular relationships observed between these samples and any other. Lastly, scale formation begins near the end of the larval period (Fuiman 2002). The formation of scales involves a number of changes to the external surface of a fish, as mineralization and collagenous deposition begin (Olson and Watabe 1980).

I found the modified version of APE (Beamish and Fournier 1981) to be useful in describing these results. The prime example of this can be seen in the EtOH stored all-ages group validation model: the R^2 (92.79), RMSECV (3.97 days), %RMSE (8.02%), and RPD (3.72) all seemed to indicate a relatively strong predictive model, similar to positive results seen in the literature (table 2.1). However, significant deviations from predicted values in early age groups were not reflected in any of these measures- it was not until I attempted to compare the ageing accuracy with traditionally derived ages that I realized how inaccurate the model was at these lower ages. As most FT-NIRS fish ageing studies to date have not found a significant, cross-taxa relationship between absolute error and age, the result of not including a measure of all-age relative prediction strength can be large. Despite having quite close %RMSE (~8.4% and ~6.7%) and RMSE (~4.1 days and ~3.3 days), the utility and actual predictive power of the all-age models and the age 24-49 models would be radically different (APE of ~34% vs. APE of ~6.7%).

Application of APE to FT-NIRS predicted ages does face some challenges, however. First, using known-age fish causes an increase in the weight any outlier predictions have on the average. As the denominator for any given fish in the equation, X_i, necessarily includes any given X_{ij} when using traditionally determined ages, the aberrant reading has a weakened impact on the overall error determined. Using known ages does not do this, and thus any outlier predictions are not self-mitigating. Another complication was found due to the impact of equally distributed error throughout the sample set. Prediction error did not significantly correlate with age (figure 2.14), meaning that the impact of error in early ages (with a lower denominator corresponding to that age) was significantly greater than that in later ages. The result is that excluding younger

fish from the analysis changed the resultant APE in a highly exaggerated manner- the set including only fish age 24-49 had an APE between ~52% - 26% less than the other sets. While there is a possibility that this is a result of the biological discontinuity between these ages (as discussed above), if similar trends persist then older ages will appear more accurate than younger ones by default. The weighting of each error in age prediction by the age itself is also, however, a primary benefit of APE. Accuracy comparisons between studies, if error continues to be unrelated to age, will need to take into account relative age classes. Despite this, I found APE a necessary metric to measure this output.

Perhaps the most useful comparison between models, and a way to get around the error discussed above, was to examine the percent of samples in a given model accurate to 10%, 15%, 20%, 25%, and 30% of true. While %RMSE does an excellent job of allowing quick comparison between model usefulness, it was found to hide a large number of inaccurate predictions in younger ages. The purpose of %RMSE is to show average error relative to the ages in the population, but by using only the maximum age it allows significant error in younger groups to go unnoticed. Conversely, looking at model accuracy as a "% within" a given range of error provides a more representative look, with error being weighted relative to each age. This allows comparison similar to APE (which also balances error by age), but in a slightly more descriptive manner since it doesn't average per sample error together in an overall index. Tests of the percentage of predictions within a given number of days of true age for the sample set was similarly useful and can be directly utilized to determine the viability of this method for the accuracy needs of a given application.

The tests of error simulating calibration ages suggest that error introduced by utilizing traditionally aged samples might be minimal in certain circumstances. When error was randomly distributed about the true ages, the APE compared to true values was quite low, remaining below 9 even with input ages which were on average 10% off. Traditional measures of model accuracy, however, were very weak in these cases (poor R^2, RMSE, and APE from input ages). Conversely, directional (positive or negative) error resulted in models which produced high measures of model success yet predicted values which were significantly farther from true ages. Had true ages not been known in these samples, randomly distributed error would have been interpreted as a predictively weak model, while those with directional error would have been deemed strong. Due to the prevalence of age-dependent biases in many ageing methodologies (i.e., older fish scales being less representative of the fish's true age, or the difficulty distinguishing otolith growth rings in some taxa past a certain age), the possible error in input ages will be essential information when interpreting FT-NIRS ageing model utility.

The differential feeding error simulations performed here, while useful as an expectation reference for erroneous inputs, lack the biological variability which would be observed in a real population. The error shown in age assignment in Jones and Brothers (1987) is the result of different feeding schedules, which causes a change in the periodicity and start of otolith formation. The biological effects of malnutrition which this indicates would likewise be expressed in other tissue, including on the skin and scales. The results of reperforming the Jones and Brothers (1987) experiment using FT-NIRS could therefore be quite different, since the spectra produced by samples in each feeding group would likely differ. In other words, while the models created here work to

show the potential of incorrect calibration age assignment, the actual results would likely be more complicated by inclusion of associated biological variability.

The potential for time and cost savings using this method are promising. While studies analyzing potential cost savings of FT-NIRS ageing have been hindered by the question of how many samples will still need to be traditionally aged for calibration, the reduction in handling time alone is significant (Robins et al. 2015). Not only is the time spent handling/ visually examining samples in FT-NIRS ageing studies significantly lower than traditional ageing methods (Robins et al. 2015; Wedding et al. 2014; Passerotti et al. 2020a), the ability to cut out the otolith removal process, as done here, suggests even more streamlining potential. While attaining otoliths is relatively fast and simple for some populations, removing otoliths from smaller taxa or younger individuals can require significant time, effort, and skill (Secor et al. 1992; Geffen 1992). Not only does using whole-fish scans take less time than removing otoliths first, it does not require the same damage to the specimen.

The success found here using whole fish justifies further investigation. The heterogenous nature of whole fish scans when compared to more homogenous structures (such as otoliths) might prevent the fine accuracy which could potentially be found elsewhere, as sample homogeneity avoids numerous issues of non-meaningful variation in spectra (Miller 2001). However, models created using appropriate calibration sample size might still, as shown here, be sufficiently accurate for many applications, while avoiding many drawbacks and costs of hard structure extraction. The higher number of loadings used for regression models here when compared to those often found in the literature might be partly explained by this- as molecular composition varies with age

across many tissue types (Vance et al. 2016), a scan of the whole fish would reasonably be expected to produce more signals which are found to correlate with age. The molecularly varied composition in biological structures in general is a consistent frustration in NIRS, which is primarily overcome through inclusion of sufficient sample size, sufficient biological variability, and appropriate model creation methods (Bobelyn 2010; Miller 2001).

The usefulness of this technique at the moment will be highly dependent upon the required accuracy for any given application. In general, these results compare favorably to other daily age studies of teleosts. Similar to Savoy and Crecco (1987) with American shad (*Alosa sapidissima),* most samples from were able to accurately placed within a 2 or 3 day window, which was viewed as a high degree of precision for use in determining hatching intensity and periodicity. The combination of high accuracy and precision found here is sufficient for many ageing studies, though daily resolution in populations which have been thoroughly validated for traditional ageing is still likely to require otolith extraction and band counts.

One exciting possibility given the success in age estimation found here is the use of *in situ* ageing. Handheld NIRS machines have been around since the 1990's and have considerably improved in both accuracy and portability in the years since (Lysaght et al. 1991; Alcalà et al. 2013). Incorporation of this technology, if the accuracy in age assignment found here holds in other taxa and age groups, would allow for rapid age estimation in the field. The number of samples which are sacrificed for ageing purposes would need only include those used to calibrate models with traditional age estimates, as scans are performed quickly enough (generally less than 60 seconds per sample) to allow

for post-scan release. Studies which target specific age groups for lethal sampling would likewise have a lower incidental mortality rate as age can be assessed in the field more accurately than by length-based estimates. The utilization of fin clips in chondrichthyan taxa for FT-NIRS ageing by Rigby et al. (2014) suggests that this implementation might provide benefits for a wide variety of ichthyofaunal taxa. While the durability of such handheld NIRS devices might need to be improved before such use, the possibility is nonetheless well worth exploration.

This chapter has shown that NIR spectra recorded from whole fish *M. saxatilis* are able to be strongly correlated with age via PLSR. The data also shows that both frozen and EtOH stored samples are capable of creating strong predictive models, and that even combined storage sets are capable of incorporating sufficient biological variability in model creation to predict ages accurately. These results show the lowest absolute RMSECV for any FT-NIRS study to date, and comprise the most well-distributed calibration set examined in the literature. The strength of the predictive models created here across calibration set frequencies was surprising and highly encouraging.

CHAPTER 3

TRADITIONALLY AGED RAW AND CLEANED FINETOOTH SHARK VERTEBRAE AND RELATIVE FT-NIRS ACCURACY

Introduction

Traditional ageing in chondrichthyans has, as previously mentioned, suffered a lag relative to the ageing of other fishes (Cailliet et al. 2006). Validation studies of the few structures which can be aged in elasmobranchs have also struggled with a number of issues specific to the taxa in question. As many elasmobranchs are long lived and difficult to raise in captivity, validation methods such as captive rearing and rear-and-release have rarely been used (Goldman et al. 2012). Instead, many shark validation studies rely upon mark-release using oxytetracycline (OTC) (Cailliet 2015). While useful, this reliance does cause a discontinuity in the confidence of vertebral increment periodicity across age classes, with older age classes possessing a much higher degree of age uncertainty (Rigby et al. 2015). This is the result of both the short-term duration of most mark-recapture studies as well as the relative paucity of samples in high age classes. Studies using OTC are also quite expensive, causing them to often have low sample numbers even in younger age classes (Cailliet 2015).

The unique life history traits of elasmobranchs make their reproduction, growth, and aging hard to study (Cailliet 2015; Harry 2018). Sampling efforts overall are often

difficult due to the size, mobility, and seasonal movement patterns found in many elasmobranchs (Cailliet et al. 1983; Cailliet et al. 2006; Cailliet 2015). All lack many of the calcified structures used for the ageing of bony fish (scales, otoliths, opercula; Cailliet et al. 1983) and many deep sea species seem to lack the banding required for conventional ageing techniques (Kyne and Simpfendorfer 2010; Cotton et al. 2011; Burke et al. 2020). In species where such banding is found, validation studies are unfortunately rare for the majority of species for which ageing has been applied and are frequently hampered by low sample sizes (Cailliet et al. 2006; Harry 2018; Burke et al. 2020). A recent review of systemic age underestimation in chondrichthyans found that, of the 58 validation studies examined, 57% were based on fewer than 10 samples and 17% used only a single individual (Harry 2018).

The utility of FT-NIRS in ageing chondrichthyan taxa has been less well explored than in bony fishes. The entirety of published papers which focus on this potential use currently consist of Rigby et al. (2014), Rigby et al. (2015), and Arrington et al. (2019). All found that a correlation between spectra and age existed, though to differing degrees in the species and structures examined. In elasmobranch spectra explored, a proposed mechanism responsible for this correlation has been hydroxyapatite $3(Ca_3PO_4)_2 \cdot Ca(OH)_2$ (Rigby et al. 2014; Rigby et al. 2015) though this has not been verified.

Table 3.1. Structures and topics explored in elasmobranch ageing with FT-NIRS analyses. %RMSE was calculated as RMSECV / maximum age. * indicates that the max age for use in the %RMSE calculation was estimated from figures.

Species	Structure	*N*	R^2	RMSECV	%RMSE	Source
Squalus megalops	Dorsal Fin Spine	97	0.82	2.41	9.64	Rigby et al. 2014
Squalus megalops	Vertebrae	97	0.89	1.85	7.4	Rigby et al. 2014
Squalus megalops	Fin clip	97	0.76	2.67	10.68	Rigby et al. 2014
Squalus montalbani	Dorsal Fin Spine	95	0.73	2.96	9.55	Rigby et al. 2014
Sphyrna mokarran	Vertebrae	80	0.83	2.48	6.34	Rigby et al. 2015
Carcharhinus sorrah	Vertebrae	102	0.78	1.23	8.98	Rigby et al. 2015
Sphyrna mokarran - verified ages	Vertebrae	76	0.89	0.87	8.52	Rigby et al. 2015
Carcharhinus sorrah -verified ages	Vertebrae	99	0.84	0.88	8.97	Rigby et al. 2015
Raja rhina	Vertebrae	648	0.88	1.41	7.83*	Arrington et al. 2019

Finetooth sharks (*Carcharhinus isodon*) are a coastal species found offshore from Florida to North Carolina. In South Carolina they are primarily found in estuaries and nearshore waters, which are used as nursery sites by the species (Castro 1993). While the population in the US is not currently considered an overfished species by most sources, they are the target of a moderately sized fishery off the Southeast US coast (Carlson et al. 2003). While they are assigned to the small coastal shark (SCS) complex by the National Marine Fisheries Service (NMFS), there exists within this category a great deal of variability in life history parameters.

Age validation for *C. isodon* has not yet been achieved for a variety of reasons. Bomb radiocarbon dating requires taxa to be longer lived than *C. isodon* or else to have samples which were collected closer to the period of ^{14}C influx (Kalish 1993). Low recapture rates limit the possibility of either physical or chemical mark-recapture analyses. Of 2773 *C. isodon* tagged since 1992, SCDNR had only received a 1.88% recapture rate as of 2019 (Vinyard et al. 2019). While the relative margin increment ratio was used to validate annual band formation in individuals 0-3 years of age, other taxa have demonstrated that this should not be considered applicable to age classes beyond this period (Conrath et al. 2002; Vinyard et al. 2019; Cailliet and Goldman 2004).

Recent evidence suggests that the populations in the Gulf of Mexico and Western North Atlantic might be distinct and therefore require specific regional management (SEDAR 2007; Vinyard et al. 2019). Among these are differences in reproduction (Driggers and Hoffmayer 2009), genetic structure (Portnoy et al. 2016), growth and size at maturity (Vinyard et al. 2019). This differentiation in populations highlights both the relative ignorance concerning a number of basic life history parameters found in many elasmobranchs and the increased need for appropriate management.

Given the difficulties of traditional age assignment in chondrichthyans, as well as their potential vulnerability to exploitation, there is a compelling interest in expanding the number of taxa for which FT-NIRS ageing has been verified (Cailliet et al. 2006; Rigby et al. 2014; Cailliet 2015; Musick 1999). While this is particularly true of species which are of commercial interest or are commonly found as bycatch, the ability of FT-NIRS to expand the number of taxa for which life history information is available is also promising (Rigby et al. 2014). Accordingly, the present study was designed to 1) expand

the number of elasmobranch taxa for which FT-NIRS ageing has been verified, 2) examine the importance of sample preparation to model accuracy, 3) explore the possibility of multi-preparation method model creation, and 4) determine potential issues of age-dependent error in the sample set used.

Methods

Sample Preparation

A set of 197 traditionally aged finetooth vertebra were scanned at $16cm^{-1}$ resolution at 64 repetitions. The specimens were initially captured between April 2002 and September 2016 (Vinyard et al. 2019). Cervical vertebrae anterior to the origin of the first dorsal fin were removed in the field. These specimens come from fish of various ages, ranging from 0 (less than a month) years old to 10 years old (with an assigned birthdate of June 1st based on historical umbilical scar data). Specimens were stored in 95% ethanol, but air dried before scanning. Initially, samples were unbleached and contained some still-present adjoining collective tissue (henceforth referred to as the "raw" samples).

Following scans, each vertebra was cleaned again using a scalpel, with adjoining tissue manually removed. Samples were then left to soak in a 1:2 sodium hypochlorite bleach-water solution for 40 minutes, before being rinsed and dried at 42º C for 2 hours. Scans were performed as described above utilizing the now dried specimens ("cleaned" samples). The raw sample set consisted of 197 individuals for which consensus ages were successfully derived from Vinyard et al. (2019). The bleached sample set consisted of 199 individuals, including 2 which were rejected from inclusion in the raw set due to

drying out at an unknown previous point. As bleaching and drying samples homogenized presentation and removed EtOH signals, it was judged that their inclusion in the bleached set was unlikely to have any negative impacts. This was later determined to be the case through each having no impact on the leave-one-out cross validation greater than would be expected from a single sample's removal.

Data Analysis

A leave-one-out analysis was performed on each sample set (raw and cleaned), as described previously, using OPUS software (version 8.2; Bruker Scientific, Billerica, MA). Each set, separately, also had validation sets (PLS-regression) created using a decreasing frequency of samples as a calibration set and the remainder being used as a test set (frequencies included 0.5, 0.33, 0.25, 0.2, and 0.1). As in the previous chapter, the assigned set for each sample was determined by sample number, without regard to age. Results were recorded in terms of R^2, RMSE, %RMSE, and RPD. The raw set utilized spectral preprocessing including a 1st derivative transformation of the 9403.7 – 6094.3 cm^{-1} wavenumber range. The bleached sample set spectra found a 1st derivative transformation in the 9400 – 7496 cm^{-1} and 6104 - 4600 cm^{-1} ranges to be the most advantageous.

To compare the degree of difference found in the wavenumbers which correlate with age in the cleaned vs. raw samples, each of the two sets were combined (n = 396). This combined set included 2 scans from each individual, including one before cleaning and one following it. A leave-one-out cross validation analysis was performed on this new set. The test-set validation analysis was performed again, this time using only raw

samples as the calibration group and only cleaned samples as the validation group. Following this, the assignments (calibration vs. test) were switched, with cleaned samples being used to calibrate the model and raw samples used to test it. Finally, a test set utilizing half of all samples from each preparation type as a calibration set and the other half as a test set was performed. Each individual was therefore included in both the calibration and test spectra, either as a cleaned or raw sample (i.e., cleaned scans from individual 1 were used to calibrate and raw scans from individual 1 were used to test, etc). As this corresponded to a calibration set frequency of 0.5, comparisons between mixed preparation and homogenous sets were able to be performed more directly.

As many of the samples used were traditionally aged to less than a year old (with many being labelled as age 0 if captured during the month of birth), the use of APE as a measure of accuracy was rejected. Instead, accuracy was calculated as the percent of samples aged to within 0.5 years, 1 year, 2 years, and 3 years of true. This value was calculated for each of the leave-one-out validation sets to allow for accuracy comparisons between the raw, cleaned, and mixed sample sets.

A length-spectra regression model was also constructed, using fork lengths obtained for each specimen at the time of collection. A leave-one-out analysis was performed to test the strength of this relationship, with results again recorded in terms of R^2, RMSE, %RMSE, and RPD. This was performed using both the cleaned and raw scan sets separate from one another.

Results

Raw Set

The raw sample set produced strong predictive models in both the leave-one-out and test set validations. RMSE ranged from 1.51 (for the leave-one-out model) to 2.19 (for the 0.1 calibration set test validation), with associated %RMSE of 6.76% to 9.81%. Most models were constructed using a rank 7 regression, though 3 models utilized either a rank 6 or rank 5 regression as the most advantageous. Correlation coefficients and predictive accuracy generally decreased with reduced calibration set size, with the exception of the 0.33 frequency calibration set being more accurate than the 0.5 frequency calibration set (R^2 = 92.59 and RMSEP = 1.64 vs. R^2 = 91.73 and RMSEP = 1.71). The RPD determined for each raw sample set age-spectra model was above 3 (Table 3.2).

Table 3.2. Raw sample set cross validation and test set results (*n* = 197).

Test	R^2	RMSE	%RMSE	Rank	Bias	RPD
All samples Raw L1O	93.49	1.51	6.76	7	0.0178	3.92
50% Calibration set	91.73	1.71	7.66	7	0.0119	3.48
33% Calibration set	92.59	1.64	7.34	6	0.104	3.68
25% Calibration set	90.9	1.77	7.93	5	-0.246	3.35
20% Calibration set	89.2	1.89	8.46	7	-0.433	3.13
10% Calibration set	85.5	2.19	9.81	6	-1.13	3.06
Fork Length L1O	99.27	26.7	2.08	9	0.208	11.7

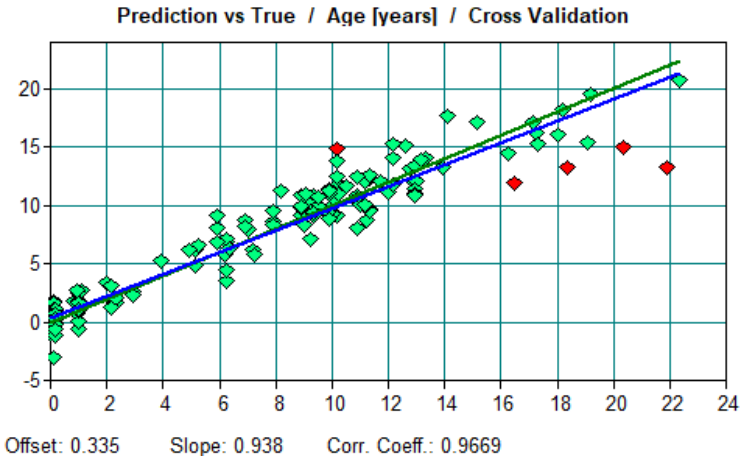
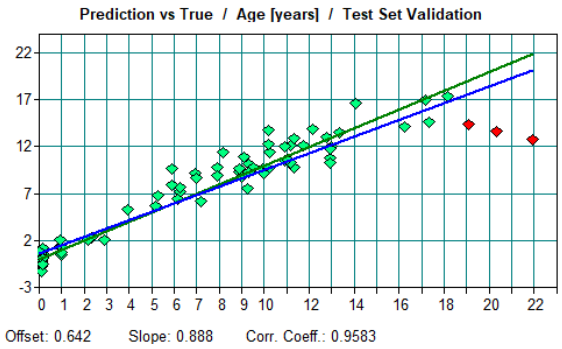


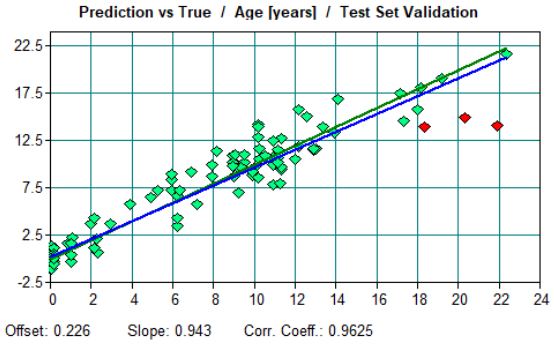
Figure 3.1. Raw sample set leave-one-out cross validation.

Most raw samples were able to be aged to within 1 year by cross validation (table 3.3). Nearly all (94.42%) were able to be aged to within 3 years of traditionally determined age. The largest absolute error found was a prediction off by over 8.6 years, in the second oldest sample in the set (true age = 21.92, predicted age = 13.22). All errors in age prediction greater than 4 years (n = 5) occurred in individuals over 10 years old, and all errors greater than 5 years (n = 3) occurred in individuals over 17 years old (figure 3.3). This error corresponded both to older age groups and an associated smaller pool of samples. Older age groups were also more likely to be underaged in the leave-one-out cross validation, showing a relationship between age and both magnitude and directionality of error (figure 3.4).

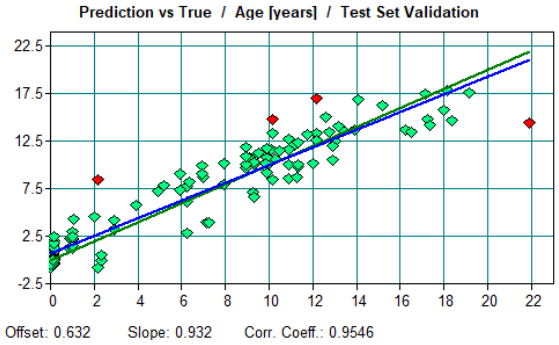
a)



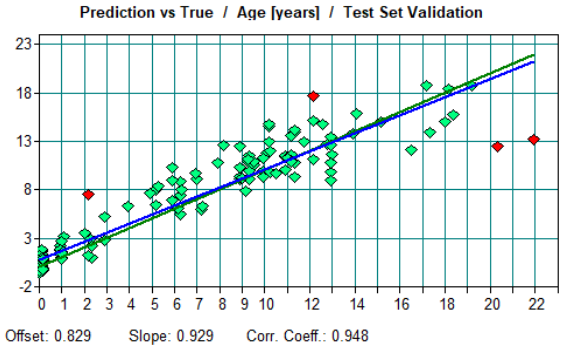
b)



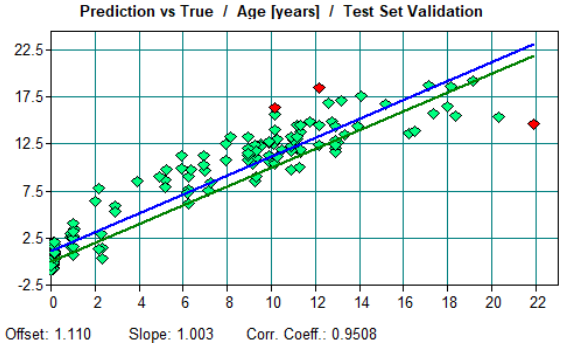
c)



d)



e)



f)

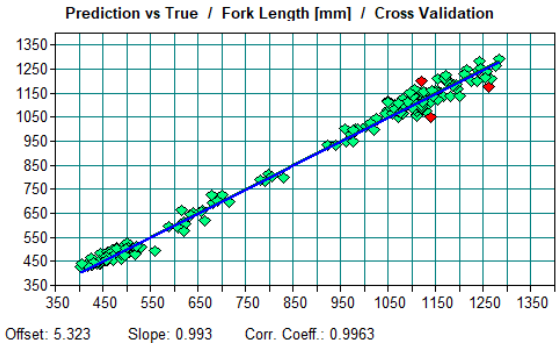


Figure 3.2. Raw sample PLSR model results for each calibration set frequency: a) 0.5 calibration set frequency (n = 98), b) 0.33 calibration set frequency (n = 66), c) 0.25 calibration set frequency (n = 49), d) 0.2 calibration set frequency (n = 39), e) 0.1 calibration set frequency (n = 20), and f) for the leave-one-out analysis of fork length (L_F). Spectral outliers are displayed in red but were not removed from any model.

To test the source of heteroscedasticity observed in figures 3 and 4, absolute error was plotted against the number of samples traditionally assigned an age to within 1 year of each sample (including itself). The increase in absolute error seen with increasing age class is consistent with issues of decreasing sample size (figure 5). All but 1 sample with fewer than 3 samples traditionally aged to within 1 year had prediction errors greater than 10%. Bartlett's test and Levine's test, however, did demonstrate that when grouped by year age classes, variance in error was still significant ($p < 0.05$).

Table 3.3. The percentage of FT-NIRS age predictions accurate to within either 0.5, 1, 2, or 3 years of traditionally assigned age. Results generated by leave-one-out cross validation of each set.

	0.5 years	1 year	2 years	3 years
% of raw set within	35.03	63.96	88.83	94.42
% of cleaned set within	43.65	68.53	85.78	93.91
% of combined set within	38.89	58.59	86.87	94.44

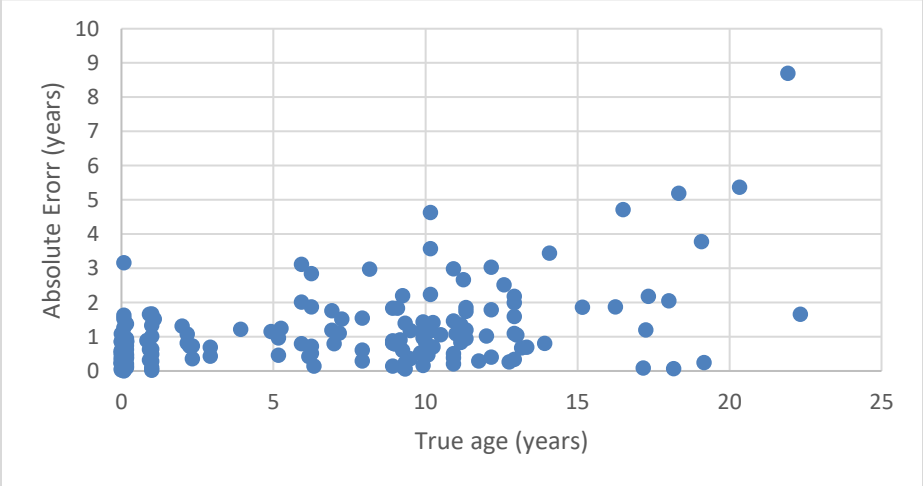


Figure 3.3. Absolute error in prediction by age of sample for raw set.

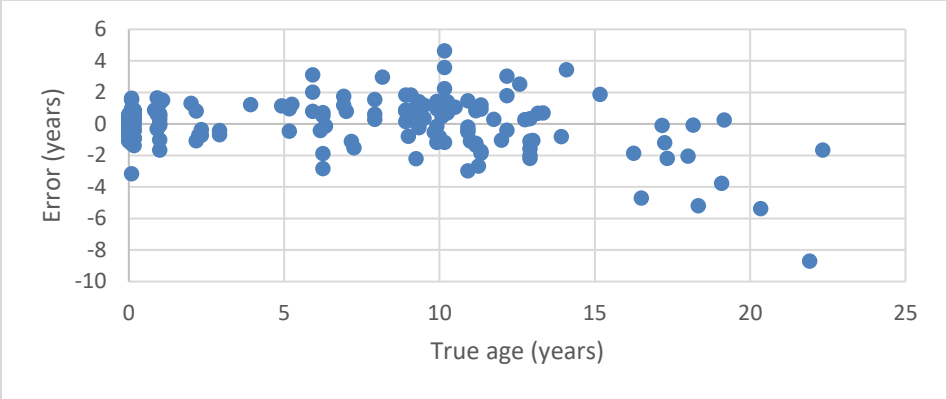


Figure 3.4. Error in prediction by age of sample for raw set.

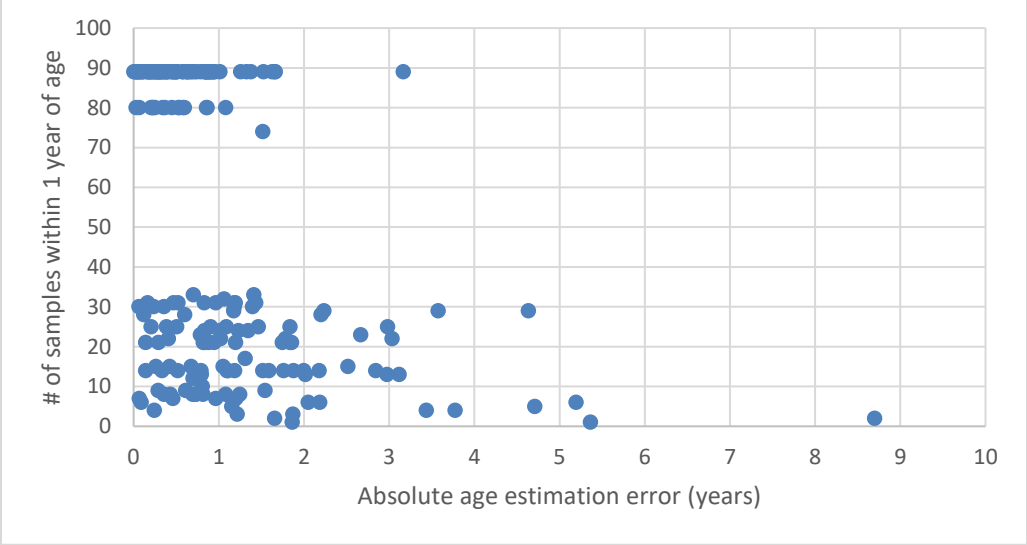


Figure 3.5. The absolute error in FT-NIRS age predictions (in years) by the number of samples assigned traditional ages within 1 year of a given sample.

The leave-one-out model using fork lengths (L_F) was highly accurate (table 3.2). The spectra - L_F regression had an RMSE of 26.7mm, which represented ~2% of the maximum sample length. The R^2 (99.27) and RPD (11.7) likewise indicate a strong spectra - length relationship.

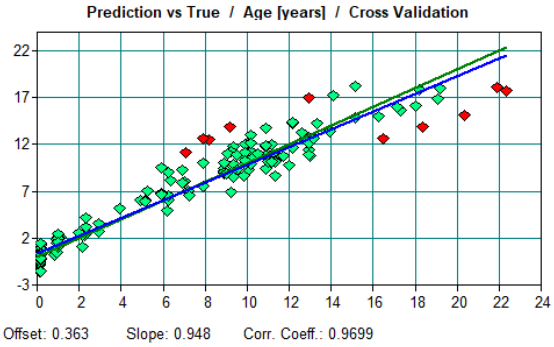
Cleaned Set

Bleached and dried samples produced models which were slightly more accurate than those of the raw set. This difference is particularly apparent in models produced using low sample frequencies- the 0.2 and 0.1 calibration set frequency models differed in R^2 (93.31/ 89.31 for cleaned, 89.2/85.5 for bleached) and %RMSE (6.72% / 8.42% for cleaned, 8.46% / 9.81 for raw). The highest accuracy (%RMSE = 5.28%) was found in the set utilizing 50% of samples for calibration, while the lowest accuracy was seen in the set which used 10% of samples for calibration. Each cleaned set model had lower RMSE than its equivalent raw set model.

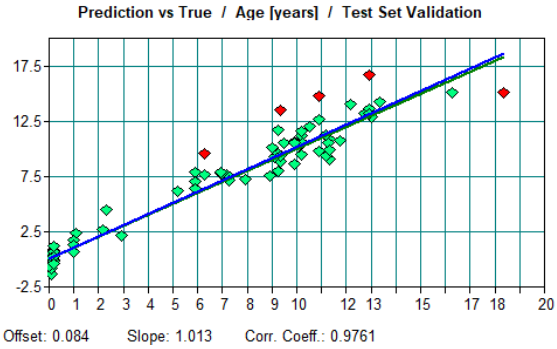
Table 3.4. Cleaned sample set cross validation and test set model results (*n* = 199)

Test	R^2	RMSE	%RMSE	Rank	Bias	RPD
All samples Cleaned L1O	94.05	1.45	6.49	6	-0.0645	4.1
Cleaned 50% Calibration set	94.81	1.18	5.28	5	-0.151	4.42
Cleaned 50% Calibration-alternate sets	92.9	1.74	7.79	7	0.261	3.8
Cleaned 33% Calibration set	93.24	1.58	7.08	6	0.0123	3.85
Cleaned 25% Calibration set	92.33	1.59	7.12	4	-0.271	3.66
Cleaned 20% Calibration set	93.31	1.5	6.72	7	0.0612	3.87
Cleaned 10% Calibration set	89.31	1.88	8.42	8	-0.536	3.19
Cleaned F_L L1O	99.11	29.3	2.29	5	-2.51	10.7

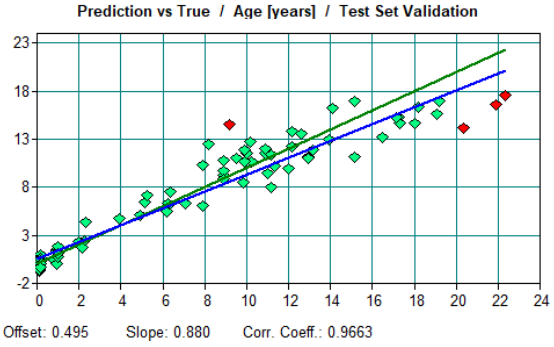
a)


Prediction vs True / Age [years] / Cross Validation
Offset: 0.363 Slope: 0.948 Corr. Coeff.: 0.9699

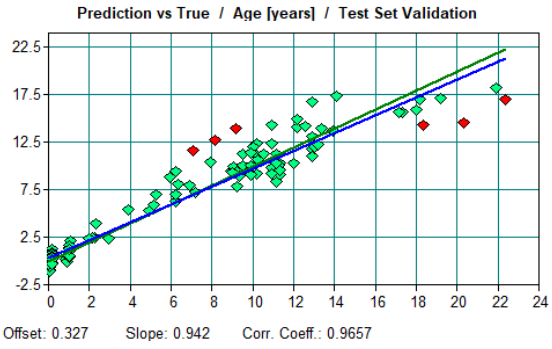
b)


Prediction vs True / Age [years] / Test Set Validation
Offset: 0.084 Slope: 1.013 Corr. Coeff.: 0.9761

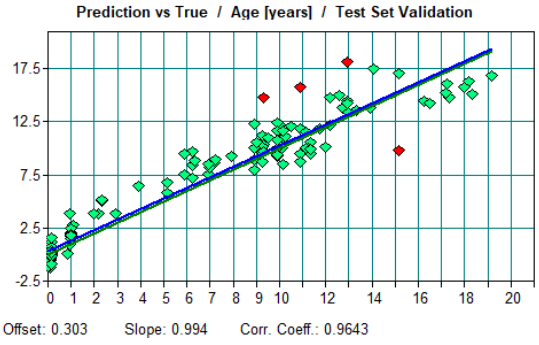
c)


Prediction vs True / Age [years] / Test Set Validation
Offset: 0.495 Slope: 0.880 Corr. Coeff.: 0.9663

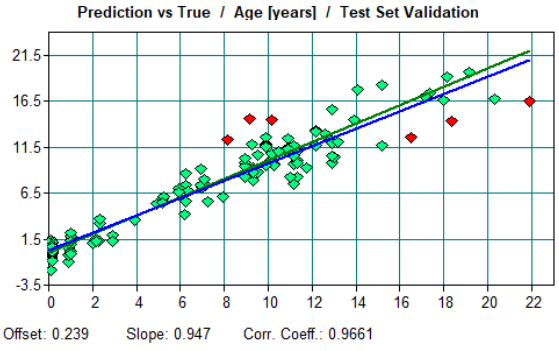
d)


Prediction vs True / Age [years] / Test Set Validation
Offset: 0.327 Slope: 0.942 Corr. Coeff.: 0.9657

e)


Prediction vs True / Age [years] / Test Set Validation
Offset: 0.303 Slope: 0.994 Corr. Coeff.: 0.9643

f)


Prediction vs True / Age [years] / Test Set Validation
Offset: 0.239 Slope: 0.947 Corr. Coeff.: 0.9661

g) h)

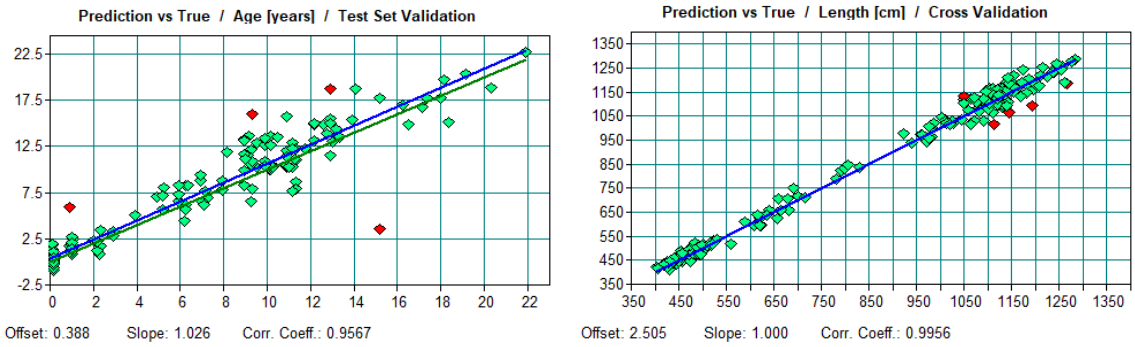


Figure 3.6. Cleaned sample set model results for all samples (n = 199): a) leave-one-out cross validation, b) 0.5 calibration set frequency, c) 0.5 calibration set frequency, sets inverted, d) 0.33 calibration set frequency, e) 0.25 calibration set frequency, f) 0.2 calibration set frequency, g) 0.1 calibration set frequency, and h) leave-one-out cross validation using fork length (L_F). Spectral outliers are displayed in red, but were not removed from any model.

Similar to the raw set cross validation, most samples were able to be aged to within 1 year of traditionally assigned age. However, a slightly smaller proportion of samples were able to be accurately aged to within 2 or 3 years when compared to the raw sample set (table 3.3). This can also be seen in figure 8, as the number of samples aged ≥4 years off of true was greater than in the raw set (n = 5 in raw set, n = 8 in cleaned set). Error directionality was similar to that seen in raw samples, with a consistent tendency to underage older samples (figure 7).

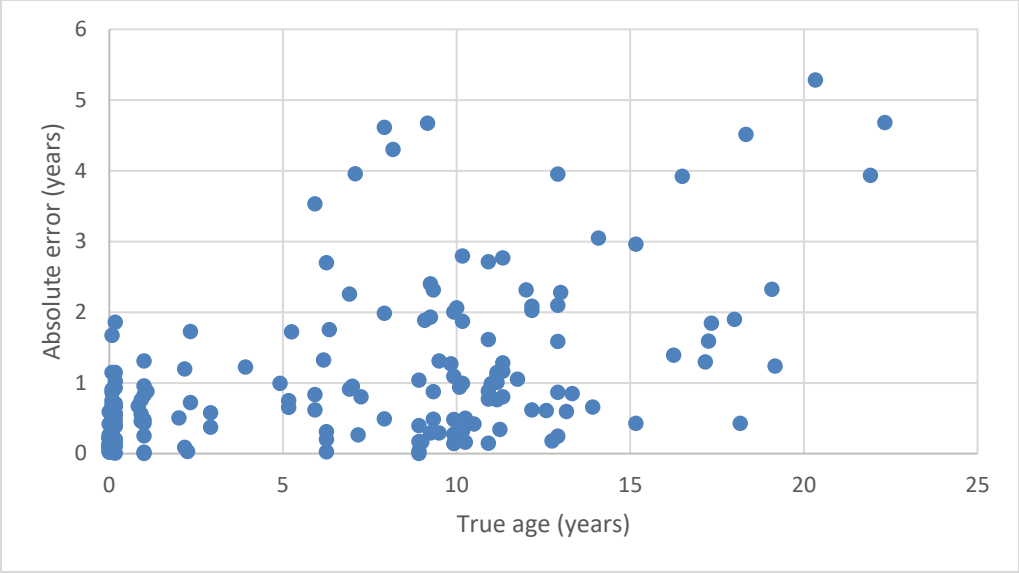


Figure 3.7. Absolute error in prediction by age of sample in cleaned set.

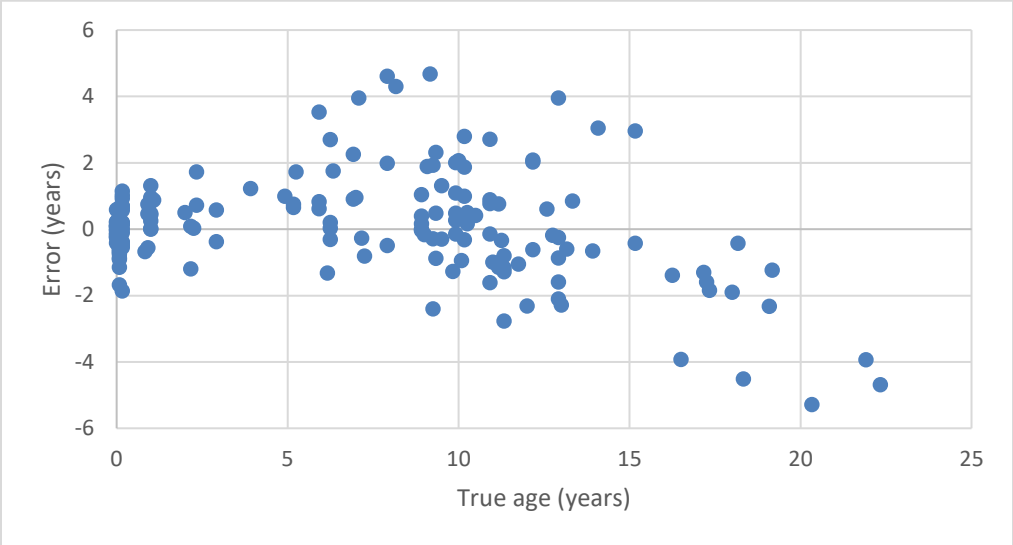


Figure 3.8. Error in prediction by age of sample in cleaned set.

Combined Set

Regressions created using this combined set utilized spectral preprocessing of a first derivative transformation in the 9400 – 5448 cm^{-1} wavenumber range. The total

combined set leave-one-out cross validation produced a strong predictive model (R^2 of 92.96 and RMSECV of 1.57), albeit at a higher rank than used for either preparation method individually (table 3.5). The test set validation which used half of each sample preparation set spectra as calibration samples was likewise fairly accurate, with an R^2 of 91.58 and a RMSEP of 1.68. The regression using only cleaned samples as the calibration set and only raw samples as the test set found no strong relationship, and had no predictive capability (R^2 = -12.6, %RMSE = 97.627, bias = -19.2). Conversely, the regression model created using the raw sample set as calibration spectra and the cleaned set as test spectra did create a weak correlation (R^2 = 78.71, RMSEP = 2.73).

Table 3.5. Combined set PLSR model results.

Test	R^2	RMSE	%RMSE	Rank	Bias	RPD
Combined L1O all samples	92.96	1.57	7.03	9	0.0237	3.77
50% Calibration set mixed sample prep	91.58	1.68	7.52	5	-0.00846	3.45
Cleaned as calibration set	-12.6	21.8	97.63	1	-19.2	0.572
Raw as calibration set	78.71	2.73	12.23	5	0.927	2.3

Similar to the raw and cleaned sets individually, most samples from the combined set were able to be placed to within 1 year of true age (58.59%; table 3.3). A comparable percentage of samples were able to be placed to within 2 and 3 years when compared to the other sets, though a small proportion were able to placed to within 0.5 years.

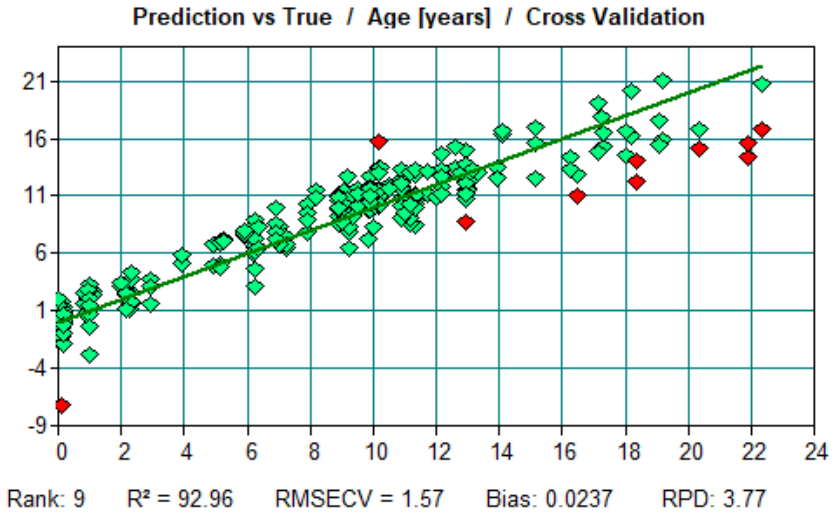


Figure 3.9. Leave-one-out cross validation using all samples (raw and cleaned).

Discussion

Accuracy in cleaned set models was highest in the 0.5 calibration frequency analysis, though this is likely an artifact of older samples being excluded from the test set by chance assignment. Switching the calibration/ test sets for this model resulted in a large change in accuracy consistent with this hypothesis. This chance assignment of older test set individuals can be seen in figure 3.6(b) and 3.6(c). While this is an easy issue to correct for, it highlights the importance of calibration sets being representative of all test samples. As older individuals are significantly less common in age analyses in general (particularly in elasmobranchs) (Cailliet et al. 2006; Harry 2018), calibration sets without sufficient samples at older age classes will likely compound any estimation errors inherent to older ages. This difference in model accuracy between the 0.5 frequency calibration group PLSR models also highlights the need for consistency of set assignment in FT-NIRS papers, as discussed in detail in the previous chapter. Many FT-NIRS ageing studies have not fully described how samples are assigned to either test or calibration

sets, which complicates any comparisons of accuracy and spectra-age correlation between studies and taxa. As seen here, while the overall correlation was strong in both models, there was a distinctive difference in metrics of model success (including a %RMSE difference of 2.5%).

The rejection of APE here was prompted by the sample age distributions. As most samples were young (between 0 – 0.167 years old), all error was highly exaggerated. The inclusion of age 0 individuals compounded this problem, as the APE formula would have a theoretical denominator of 0 in these cases. Solutions were explored, such as using 1 as the standard denominator for each sample traditionally aged to under 1 year old or adding a constant to all age assignments, but this change would have resulted in APE which was atypical to other studies. While weighting error by sample age is a useful metric of determining accuracy, and arguably a strong point of using APE overall, in this case it prevented strong comparisons between studies, which was the primary purpose of APE in the other chapters. Instead, the percentage of samples aged to within 0.5, 1, 2, and 3 years of traditionally determined ages was deemed more informative. While this issue is unlikely to be present to the same extent in other sample sets, where an age of 0 is an unlikely assignment, the magnitude of error found at low age classes here would nonetheless bias FT-NIRS and non-NIRS ageing study accuracy comparisons.

The strength of the leave-one-out cross validation model comparing spectra and length was marginally higher than has been found for other elasmobranchs in FT-NIRS studies. Rigby et al. (2014) found a R^2 of 0.81 and a %RMSE of 2.5% when using a spectral model created with *A. pallidus* vertebra with a narrower band of lengths, compared to the R^2 of 99.27 and %RMSE of 2.08% found for *C. isodon* here with raw

samples and R^2 of 99.11 and %RMSE of 2.29% with cleaned samples. Slightly closer was the PLSR model correlating vertebrae from *S. megalops* and length, with an R^2 of 0.94 and a %RMSE of 2.66%. The strength of the raw sample model produced here was originally thought to be caused by the presence of adjoining tissue allowing for more overfitting, given the high number of loadings used. However, model strength was still high with only 2 loadings used (R^2 = 96.56). The cleaned sample set model was likewise strong enough to dismiss this hypothesis while utilizing only 5 loadings. It is possible that the difference in length resolution observed between *C. isodon* and *A. pallidus* is due to life history traits, but without further elasmobranch taxa verified with FT-NIRS this is not yet possible to pursue. The increased strength of FT-NIRS models created using length when compared to age is possibly due to the decreased uncertainty of measurement (Rigby et al. 2014). If input ages are less accurate than input lengths, then the resultant models would show the pattern found here and in previous studies (e.g. Rigby et al. 2014).

Overall ageing accuracy found here compares very favorably to what has been found previously in elasmobranchs (table 3.1). While the cleaned set RMSECV of 1.45 years was not the lowest found for elasmobranchs using FT-NIRS (0.87 years in verified age *S. mokarran*, Rigby et al. 2015), it is comparable to the RMSECV range of 1.23 – 2.96 years found in other elasmobranchs when using whole sample sets (Rigby et al. 2014; Rigby et al. 2015). These results verify the use of FT-NIRS for age determination in *C. isodon*. The overall RMSECV and %RMSE found so far in chondrichthyans suggest that FT-NIRS has just as much resolving power for cartilaginous fishes as for bony fishes.

While this chapter uses traditionally determined ages and true ages interchangeably, it is important to note that these samples were initially aged using the band count method rather than from a direct validation approach. As with other FT-NIRS ageing studies, this means any error included in these age assignments was worked into the PLSR models in both calibration and test sets. The higher error seen here in older age predictions might be at least partially explained by this, as traditional ageing likewise sees higher error with older samples (Rigby et al. 2014; Rigby et al. 2015; Harry 2018). Samples for which no consensus age could be traditionally reached were excluded from the beginning, as no appropriate input age could be determined (Vinyard et al. 2019). While the results of the previous chapter suggest that minor input errors are unlikely to unduly influence model success, the directional error often found in older elasmobranchs is exactly the type which was found to exert a more powerful impact on metrics of predictive capability.

The statistical determination that standard deviation of error varied by age class through Bartlett's test was a helpful and unique approach. While the degree of error is often viewed as increasing by age class when ageing elasmobranchs, the statistical weighting of nonparametric sample sets is often performed inappropriately (Cailliet et al. 2006). It was determined here that while sample sizes varied dramatically between age classes, this failed to entirely explain the increase in absolute error seen in older groups. This was confirmed by Levine's test, which is arguably better suited to such datasets. However, the introduction of biological variability is not itself represented in this statistic; a larger number of samples in older age classes (some of which included in the performed Bartlett's test had only 2 samples traditionally assigned to within 1 year of

age) might change this conclusion by both redefining the spectra - age regression itself and by directly changing the variability per age group. A total of 6 single-year age classes were excluded from the test entirely, as they consisted of only a single sample within that year. This included two of the highest age classes in the sample set (including 20 – 21 years old and 21 – 22 years old). While our sample size was considered high when compared to many chondrichthyan ageing studies (Harry 2018), it was arguably insufficiently well distributed to provide the necessary variability required for robust model creation.

The higher inaccuracy of models which combined the two preparation methods was not unexpected. However, the models which included either all spectra (the leave-one-out cross validation) or an equal number of both preparation types were still capable of predicting age to within ~7.5% of maximum. They also resulted in models which were strong on most measures of model success, including R^2 (92.96 and 91.58), bias (0.0237 and -0.00846), and RPD (3.77 and 3.45). This suggests, similar to what has been mentioned previously, that inclusion of sufficient variability in the calibration set can create a robust model, though with accuracy decreasing with the greater variability incorporated (Bobelyn et al. 2010). A model using mixed preparation sample spectra might therefore be viable, if not ideal, at determining ages. Most samples were able to be placed within 1 year of traditionally assigned ages, and the placement of samples within 2 or 3 years of true age was similar to that found in the raw-only and cleaned-only sets. Given the low sample sizes commonly seen in chondrichthyan ageing studies and the variable methods of preparation and storage used (discussed more thoroughly in chapter 1), this might be quite relevant for calibration set creation. Future studies examining the

impact of storage age of elasmobranch vertebrae will likely help resolve this possibility further; the ability to create a model using calibration samples from historic datasets would dramatically increase the representation possible for each age class, but this would only be useful if the consequent decrease in accuracy from incorporating different storage time and preparation methods is found to be smaller.

This does not hold for models using a different preparation type for calibration and test sets, however. The model which used only cleaned samples for calibration found no significant relationship between age and spectra whatsoever and had no predictive capability (%RMSE = 97.627). Conversely, the model which used only raw samples for calibration was still able to create a correlation, if only weakly (R^2 = 78.71, %RMSE = 12.226, bias = 0.927, RPD = 2.3).

This difference in performance between these models was unexpected and might be the result of spectral preprocessing. From figure 3.10 it can be seen that the regression coefficients used in model creation varied dramatically by sample preparation, with the least amount of noise and the least number of loadings used in the cleaned sample set. This might indicate that the spectral signals necessary to obtain the strong age-spectra correlation found in the raw-only sample models were not found in the cleaned sample set, or that they were not found to correlate with age in the same manner. Unsurprisingly, this suggests that homogenous distribution between calibration and test sets is of vital importance to model accuracy. The ability of raw sample spectra to predict cleaned sample ages, however, suggests that the importance of this requirement will vary widely depending on the spectral signals used by a given model.

Another way to explain this difference in predictive power of the raw set when used as calibration for the cleaned test set is found in the more heterogenous nature of the raw samples. The presence of adhering, non-calcified tissue in the raw set samples necessarily introduces a greater variability of molecular bonds which can be detected with NIRS. The presence of a subset of the bonds used for the raw sample calibrated PLSR model within the cleaned sample spectra would therefore allow for some degree of correlation to be detected, such as that which was seen. The more homogenous nature of the cleaned tissue, on the other hand, allows for a model which utilizes fewer spectral signals (seen in the lower number of loadings used in model creation). The result of this narrower spectral correlation model is that it would be unable to predict ages in samples which did not include these more specific criteria. This can also be seen in figure 10, where the relatively few peaks and little noise of the cleaned regression coefficient would require more specific sample composition to create a good fit.

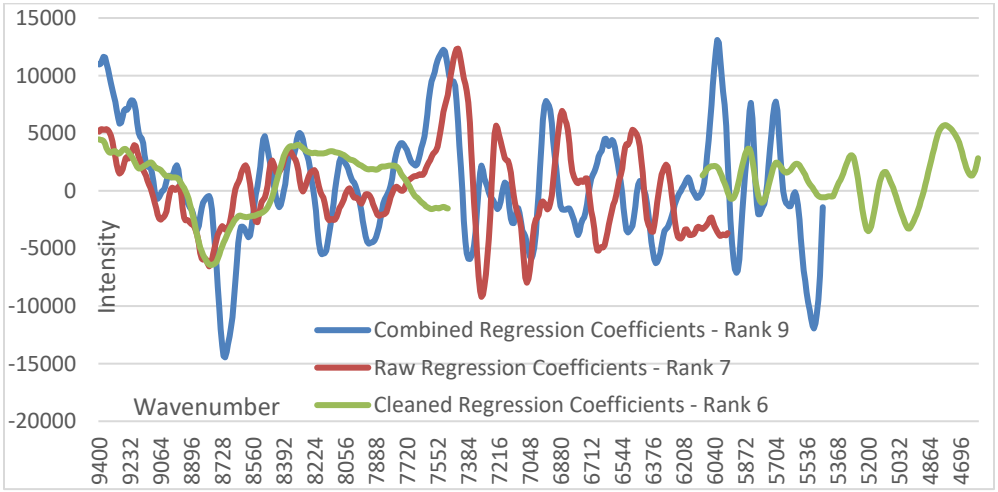


Figure 3.10. Regression coefficients for each of the leave-one-out models produced (raw samples only, cleaned samples only, and combined samples). Blank areas for a given line indicate that those wavenumbers were not used in model creation.

The similarity in success found between raw and cleaned sets was surprising. As sample homogeneity is a large component in NIRS model success, it was hypothesized that the raw sample, with its associated non-calcified tissue, would be significantly less predictive (Miller 2001). Contrary to expectation, the models produced using cleaned samples were in most cases only marginally better in both predictive power and metrics of model strength than those using raw samples. As the extra time and preparation required for sample cleaning, bleaching, and drying can be significant when working with large sample sizes, the ability to remove these processes without an accompanying loss of accuracy would clearly be desirable.

There are a number of possibilities for this similarity in model success. First, the higher number of loadings used in the raw sample models introduces the possibility that not only calcified tissue was found to correlate with age. Successful use of whole larvae in the previous chapter and elasmobranch fin clips in previous NIRS ageing studies (Rigby et al. 2014) reinforces the possibility that many tissue types are likely to produce molecular spectra which can be correlated with age. Conversely, the increased number of loadings could be the simple result of model overfitting, which would likewise benefit from the wider array of molecular bonds present in the raw samples. That the raw leave-one-out cross validation model maintains an R^2 of 84.2 and a RMSECV of 2.35 with only 2 loadings serves as strong evidence against this explanation of overfitting.

Elasmobranchs might serve as a taxon where the use of %RMSE (Passerrotti et al. 2020) might not always be particularly helpful. As only 11 of the 197 samples in the raw sample set were over even 17 years of age, the use of the maximum age (22.33 years old) as a denominator in the %RMSE calculation increased apparent accuracy in a way not

representative of the set. Even with the higher samples having significantly higher error (discussed above), their presence served to make the set as a whole appear more accurate rather than less. The introduction of %RMSE as a way to quickly compare model success between studies was a strong step forward (Passerotti et al. 2020a; Passerotti et al. 2020b), but care must be taken before it is used as a more serious indicator of relative model success. Inclusion of a single 30 year old individual, for instance, would raise the %RMSE significantly even if it possessed a 90% error in age estimation. While APE might be an appropriate metric in other sample sets to resolve this problem, for reasons discussed above percentage of samples within *x* years of true was considered a better metric here.

The significantly higher error in age prediction found in older individuals matches what has been found previously by Rigby et al. (2015). In that study, individuals over 10 years of age were all excluded due to being identified as outliers in model calibration. While it is suggested that this might be due to having relatively few samples available at higher age classes, our analyses preliminarily suggest that this impact might not be wholly related to low sample sizes. The subsequent study (Rigby et al. 2014), however, found that individuals were able to be successfully aged up to the maximum ages of 25 and 31 years old. The higher inaccuracy of NIRS generated age predictions found here and by Rigby et al. (2015) might simply be the result of the aforementioned increased error in traditionally determining ages of older individuals. Analyzing the use of FT-NIRS on longnose skates (*R. rhina*), Arrington et al. (2019) found older individuals to be significantly underaged in NIRS cross validation and test models. Validation of a set of samples which encompasses the entire age range of a species, perhaps through bomb

radiocarbon dating, would significantly help resolving the cause of this differential age-related prediction accuracy.

The error commonly associated with elasmobranchs in higher age groups is often attributed, in part, to the slower growth found in older individuals. As maximum size is approached, tissue which is not reabsorbed is no longer continually added to, resulting in a lack of consistent banding in calcified structures such as vertebrae (Cailliet et al. 1986). This might also imply that FT-NIRS would be unable to accurately age individuals nearing maximum size, as the molecular bonds which are correlated with age would cease being added to such tissue. It is also possible that using "raw" sets like those utilized here might partially overcome this barrier; if vertebrae are no longer growing, it is the surrounding tissue which has the possibility of being reabsorbed and reformed which might carry molecular bond signatures associated with age. This research, however, is far beyond the scope of the current project, and any yet attempted. A large number of individuals who have reached maximum size and continued to age would be required (and who are able to be age validated in spite of this), which is uncommon in elasmobranch studies (Rigby et al. 2014). While the use of captively reared chondrichthyans comes with its own set of caveats concerning the comparability of growth rates to wild individuals (Van Dykhuizen and Mollet 1992; Cailliet and Goldman 2004), in this case aquaria- raised individuals might offer a unique possibility. Shorter lived elasmobranchs who have died of natural mortality in captivity could potentially be used to test the possibility of molecular bonds continuing to correlate with age after apparent cessation of somatic growth.

This chapter has successfully demonstrated that vertebrae prepared in different ways from *C. isodon* can be used to create robust and accurate FT-NIRS ageing models. While slight differences were seen in model accuracy between the raw and cleaned datasets, both were considered successful models, and the preparation type used in future studies will likely depend more upon the population being examined than FT-NIRS specific cleaning requirements. While less accurate, models which made use of samples regardless of preparation type were still capable of age prediction, which opens up numerous possibilities for the utilization of historic sample sets in model calibration. Finally, these results suggest, preliminarily, that the increased error seen in age prediction for higher age classes is not purely due to the lowered sample size available. While significantly more research will need to be done on the topic, this reinforces the abovementioned results which found that for certain elasmobranch taxa, older age classes were not able to be accurately predicted with FT-NIRS.

CHAPTER 4.

USE OF WINTER SKATE VERTEBRAL SECTIONS TO EXPLORE IMPACT OF STORAGE AND PREPARATION ON FT-NIRS AGEING

Introduction

Skates comprise the second most speciose group in Chondrichthyes, yet have seen relatively little research interest until recently, primarily due to their lack of commercial interest (Ebert and Compagno 2007; Sulikowski et al. 2003; Frisk et al. 2019). The recent increase in commercial landings, and subsequent management, has highlighted several deficiencies in our knowledge of basic life history traits across skate taxa (Elliot et al. 2020; Sulikowski et al. 2003; Sulikowski et al. 2005; Kelly and Hanson 2013). Given the critical role of age assessment in both management and ichthyological research, increasing attempts have been made to develop robust ageing protocols since the early 2000's (Goldman et al. 2012; Sulikowski et al. 2003; Sulikowski et al. 2005; Francis et al. 2001).

As with other elasmobranchs, the number of skate species which have had ageing techniques validated is relatively small (Goldman et al. 2012; Cailliet 2015). Bomb-radiocarbon dating validation has found that band-pair deposition is annual in some species (e.g. winter skates, *Leucoraja ocellata*, up to 19 years of age, McPhie and Campana 2009; Carbonara et al. 2020), though the regularity of this deposition across age

classes seems variable in other taxa (Pierce and Bennett 2009; Natanson 1993; James 2020). Marginal increment analyses (MIA) have supported annual deposition of band pairs in some cases (Sulikowski et al. 2003), as has oxytetracycline (OTC) injection (Holden and Vince 1973; Abdel-Aziz 1992; Cicia et al. 2009). Similar to many other elasmobranchs, age validation has rarely been performed equally throughout life stages, and evidence exists that some skate species cease annual band pair deposition after reproductive maturity (Natanson 1993; James 2020). Given the high risk of exploitation seen in many batoids, exploring alternate ageing strategies is of high management interest (Cicia et al. 2009; Elliott et al. 2020).

Many investigations into the suitability of caudal thorns as an age-determining structure have found mixed results. Gallaghar and Nolan (1999) found them suitable for age determination, as did Matta and Gunderson (2007) and Serra-Pereira et al. (2008), though many others noted a distinct difference in ages determined from thorns and vertebrae (Davis et al. 2007; James et al. 2014; Winton et al. 2014). However, vertebrae remain the most commonly used structure for ageing skates, and the majority of historical sample sets consist of sectioned and mounted vertebrae (Sulikowski et al. 2003; Cicia et al. 2009). The ability to utilize this group of samples for the creation of FT-NIRS calibration models could, as mentioned in the previous chapter regarding finetooth sharks, substantially improve the resolution available from this new method.

The winter skate (*L. ocellata*) is a large benthic batoid found off the coast from Canada to North Carolina (McEachran and Musick 1975; Kulka et al. 2009). They are an ecologically important generalist predator and comprise a substantial portion of the soft bottom ichthyofaunal biomass (Frisk and Miller 2006; Kelly and Hanson 2013).

Recently, differences observed in life history traits between populations have suggested reclassification of a potentially cryptic group within a nominative *L. ocellata* population (Kelly and Hanson 2013). Population dynamics of this species are little known (Sulikowski et al. 2005; Frisk et al. 2006; Kelly and Hanson 2013). Observations published within the last 3 years show the first evidence of *L. ocellata* being a highly motile species (Frisk et al. 2019), which further emphasizes the amount of basic biological information which is unknown for the winter skate. In 2003, MIA supported the assumption of annual band-pair deposition in *L. ocellata*, which was reinforced by bomb-radiocarbon dating in 2009 (Sulikowski et al. 2003; McPhie and Campana 2009). Given its status as a strongly age-validated species with significant management importance and a large number of historic samples, it was deemed a strong choice for FT-NIRS exploration.

The primary objective of the present study was to validate the utility of FT-NIRS on a mounted and stored sample set. Previous work has found NIRS ageing an accurate tool in other skate taxa (Arrington et al. 2019), as well as in other chondrichthyans (Rigby et al. 2014; Rigby et al. 2015; previous chapter), which suggests that a correlation between spectra and age should likely be found within the nominally similar vertebral tissue of *L. ocellata*. As a result, this species was considered a good option to investigate the impact of storage and preparation upon age-spectra relationships.

Methods

Sample procurement

Winter skate samples, mounted upon slides, were obtained from Sulikowski et al. (2003). Samples were originally captured by otter trawl between 1999 and 2001 off the coast of New Hampshire. Of the 304 captured, 230 had vertebrae removed from above the abdominal cavity which were subsequently frozen. Excess tissue was removed manually and samples were air dried before sagittal sectioning was performed with a mini-saw rotary tool. Sections were then mounted on a slide and polished using progressively finer sandpaper. Initial band-count ageing results can be found in Sulikowski et al. (2003), meaning that these samples have been mounted for approximately 20 years. Traditional age determination placed the samples between 0 and 19 years of age.

Due to the age of the slides, a number of samples were excluded from all analyses. Most samples excluded were a result of the adhesion failing over time. When samples fell off the slides, they likewise lost the labels which related to the determination age published in 2003. All samples which could not be definitively linked to a previous age reading were discarded, as were samples which appeared visually damaged when compared alongside the others. The result of these removals created the "All Samples" sample set (n = 168). Samples which were visually aberrant due to either yellowing or blackening were removed for the "Refined" sample set (n = 133).

Scanning

Scans were performed using a Bruker Matrix-I Near Infrared Spectrometer with a 22-mm diameter sample window, as described in previous chapters. Initially, a Teflon disk with a 2 mm aperture (Passerotti et al. 2020b) was used to ensure that light scattering played a smaller role in spectra acquisition by allowing only tissue in the very center of the sample to be exposed to the sample window. A gold transflectance stamp was placed over each sample slide to prevent light penetration from likewise confounding the results. A total of 64 scans were performed for each sample (n = 168) at 16cm^{-1} resolution, with each scan for a given sample being averaged together to create a spectrograph representative of its molecular bonds.

After initial scans were performed, the Teflon aperture was modified. The opening was expanded from a 2 mm circle to an oval with a maximum width of 9mm. This was done to allow the entirety of each tissue sample to be scanned, while still excluding as much of the surrounding resin matrix as possible. The modification of shape in addition to aperture size was a result of the distinctive "butterfly" shape of vertebral sections. Scans were repeated as described above for each of the sample sets.

Data Analysis

A PLS regression model was created for each set of scans (2 mm aperture and 9mm aperture) using OPUS software (version 8.2; Bruker Scientific, Billerica, MA). Data were optimized using vector normalization in the 9400 – 7496 cm^{-1} and 6104 – 5448 cm^{-1} ranges for scans taken with the 9 mm aperture and a first derivative transformation in the 9400 – 6096 cm^{-1} range for scans using the 2 mm aperture. A leave-

one-out validation was performed for each, as has been previously described (one sample from each set is removed, a regression is created using the remaining samples, and then the removed sample is plotted along this regression with its deviation from prediction recorded). After initial leave-one-out analysis, strong outliers were removed from the sample sets and the tests were performed again. Finally, all samples with any possible problems which might reasonably be expected to affect spectral results were removed and the leave-one-out analysis was done on the remaining sets from each aperture diameter (*n* = 133). The R^2, RMSE, %RMSE and RPD were recorded for each of these for comparison.

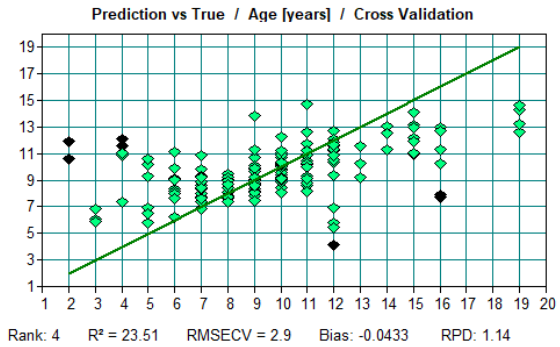
Results

No strong correlation was made between spectra and age for any sample set using the 2 mm aperture (R^2 from 23.51 to 36.79). The RPD, APE, and %RMSE likewise indicate no strong relationship between spectra and age for this set. While improvement was seen from the exclusion of initial outliers, the resultant model still did not successfully relate age to spectra. The exclusion of visually aberrant samples likewise caused modest improvement in model accuracy, but the resulting R^2 (35.89), APE (25.51), and %RMSE (12.74) were indicative of no strong predictive capability.

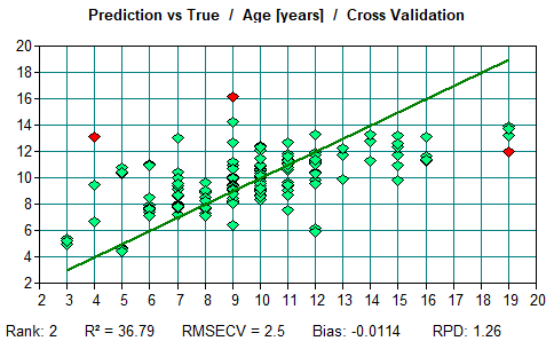
Table 4.1. Results of leave-one-out analyses using scans taken with the 2 mm aperture.

Test	*n*	R2	RMSECV	%RMSE	APE	Rank	Bias	RPD
All samples	168	23.51	2.9	15.26	29.51	4	-0.0433	1.14
Excluding outliers only	162	36.79	2.5	13.16	21.56	2	-0.0114	1.26
Excluding all problematic	133	35.89	2.42	12.74	25.51	8	0.00343	1.25

a)



b)



c)

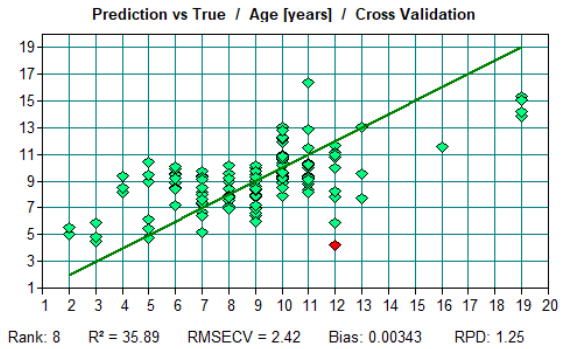


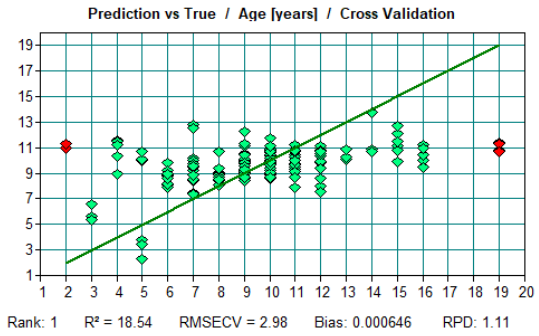
Figure 4.1. Results of the leave-one-out analysis utilizing the 2 mm aperture with a) all samples, b) exclusion of initial outliers, and c) the refined dataset.

Models created from scans taken with the 9 mm oval aperture likewise showed no strong spectra- age correlation. Most R^2 values were slightly lower than was seen with the 2 mm aperture models, with accompanying increases in RMSECV (table 4.2). The youngest and oldest samples were consistently found to be outliers, though the lack of strong age-spectra correlation naturally predicts this. Calibration set frequency tests were not performed due to the lack of significant correlation found in cross-validation.

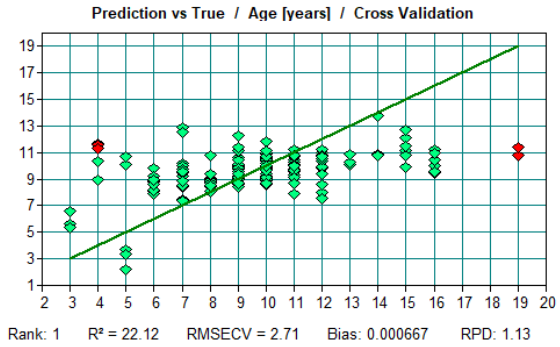
Table 4.2. Results of leave-one-out analyses using scans taken with the 9 mm aperture.

Test	*n*	R2	RMSECV	%RMSE	APE	Rank	Bias	RPD
All samples	167	18.54	2.98	15.68	30.61	1	0.0006	1.11
Excluding outliers only	163	22.12	2.71	14.26	25.32	1	0.0007	1.13
Excluding all problematic	133	24.16	2.63	13.84	28.00	6	0.0131	1.15

a)



b)



c)

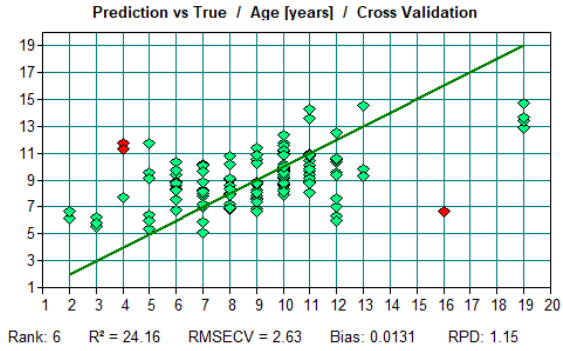


Figure 4.2. Results of the leave-one-out analysis utilizing the 9 mm oval aperture with a) all samples, b) exclusion of initial outliers, and c) the refined dataset.

Discussion

The failure of the 2 mm aperture scans to establish a strong predictive model was predicted, due primarily to the aperture itself. As vertebral centra growth occurs via areolar mineralization, focusing scans entirely on the center area of sagittally sectioned vertebrae might not allow for chemical differentiation between tissue by age (Dean et al. 2015). Chondrichthyan vertebral tissue is not reabsorbed through the life of a fish, so material which is correlated with age would only be found on the outer edges of these sections (Dean et al. 2009; Whitledge 2017). Focusing entirely upon the center of a given vertebra might therefore not include the areas which are necessary to differentiate

samples. This would also help explain the differences in wavenumbers utilized for the regression analysis- as the areas which are chemically relevant to age on the vertebrae are on the outer edge, the scans using the 2 mm aperture were unlikely to include the same relevant wavenumbers. Instead, the PLSR model would attempt to "fit" another spectral signature to input ages, which clearly did not correlate well with age.

During scanning it was noted that light visibly scattered more than was seen in other structures. A gold transflectance stamp was used to minimize background spectral interference, but the thickness of the slide itself could not be entirely covered by this. The teflon apertures placed over the laser were used to focus light more consistently upon the center of each sample tissue, but the epoxy and glass into which the samples were set nonetheless resulted in some light scattering to the sides. The observed weak correlation with spectra and age are also what would be expected from a strong, inconsistent background interference.

Another source of the background interference might be the different visible clouding effects found in the mounting media between samples. Even on the same slide the degree of this visual obstruction varied significantly (figure 4.3). The areas which were most likely to be obscured by the mounting media were the edges, which are also the areas which are most critical for creating an age- spectra regression. Even without any extraneous molecular signatures (discussed more below), the simple physical obstruction seen in the slide samples might have prevented any model success by itself. The interaction of physical light and NIR spectra is a complicated phenomenon, but blocking NIR light from reaching relevant tissue would naturally prevent molecular bonds in that tissue from being recorded (Williams and Norris 2001). Without supporting either of

these possible sources of background signals, the “messiness” and noise of the preprocessed spectra supports the idea that scans were subject to significant variability (figure 4.4).

Figure 4.3. Examples of the vertebral sections used for scanning. On the left the amount of clouding is highly variable, even on the same slide, while the right shows one of the more severely yellowed tissue samples used (which was discarded for the “refined dataset” analyses).

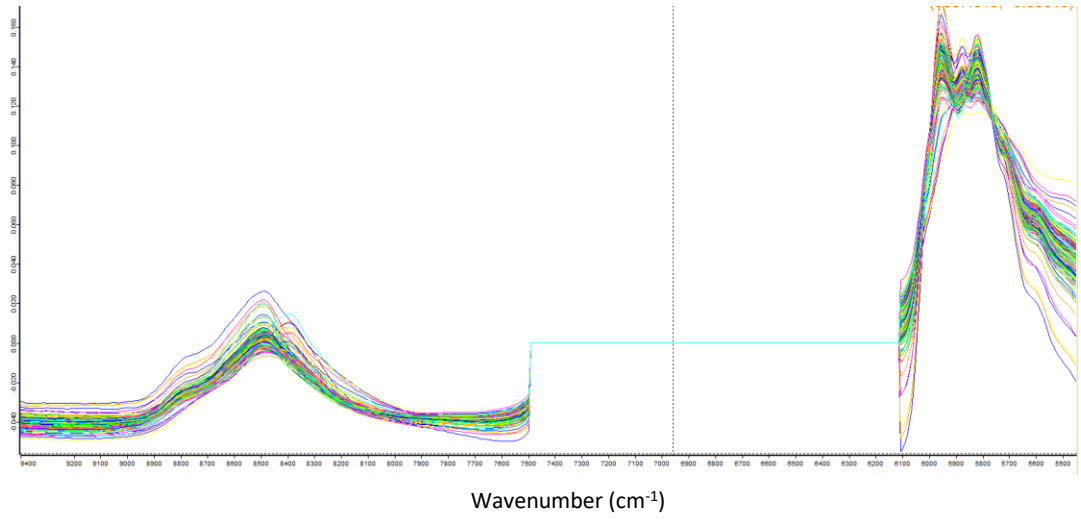


Figure 4.4. Processed spectra used to create the PLSR model for scans taken with the 9 mm aperture.

The presence of the epoxy mounting media might also introduce a confounding vibrational frequency, detectable using FT-NIRS. While excluding samples which were

visibly yellowed did not result in a substantial improvement of ageing accuracy, other less visible factors might have also influenced the spectra. The chemical structure of epoxy resin depends on the combination of constituent materials used (e.g., resin and hardener) (Ellis 1993). Any slight deviations in the mixing of these parts between samples would introduce a distinct absorbance frequency unrelated to the underlying tissue composition which would vary between each sectioned vertebra. The ageing of resin also results in a wide number of chemical changes, potentially resulting in a distinct chemical difference between samples depending on the time since they were embedded (Ellis 1993). Other studies which have made use of resin embedded vertebrae for chemical analysis incorporate a method of chemically disentangling the epoxy itself, such as the addition of a unique chemical signature (e.g. indium; McMillan et al. 2017) or by deciding, *a priori*, to analyze only elemental concentrations which have been shown to relate directly to environmental conditions of the taxa in question (Feitosa et al. 2020). In theory, consistently introduced chemical signals which do not correlate with input ages would be excluded from PLSR model creation, but if there were any inconsistencies between samples then the effect would mimic that of background interference. While embedding and mounting has a number of previously discussed benefits in band-count ageing contexts, any process which contaminates chemical structure should clearly be avoided in FT-NIRS analyses.

The chemical changes introduced by the epoxy might have affected the vertebral tissue itself as well as introducing its own contaminating presence. The curing of epoxy is an exothermic reaction, which can cause significant temperature fluctuations in surrounding tissue (Ellis 1993; Jolivet et al. 2013). Temperature has a distinctive effect

on the chemical structure of organic tissue, which could potentially break down the molecular bonds which are found to spectrally correlate with age (Disspain et al. 2016). As the heat generated by resin curing is based in part upon the volume of resin present and the ratio of hardener to resin, any slight differences in preparation and mounting could introduce chemical signatures which would vary by sample (Ellis 1993). The process of embedding otoliths in epoxy has been found to chemically alter the organic matrix of a sample, possibly through epoxy infiltrating the organic structure (Jolivet et al. 2013). As elasmobranch vertebrae are comprised primarily of hydroxyapatite within an organic matrix, a similar process could be at work here (Urist 1961).

The edges of skate vertebrae, where appositional growth occurs, differ from internal regions in a number of ways which would encourage such resin infiltration. Proteoglycan content is higher in these edge regions, though with some differences seen depending upon time of collection (Gelsleichter 1998). The presence of proteoglycan has been shown to inhibit hydroxyapatite formation in outside taxa, and a negative correlation has been found between the amount of proteoglycan found in tissue and its degree of calcification (Kemp 1984). In skates, the area just inside of the outer envelope houses an unmineralized organic matrix, as well as randomly positioned canals leading into deeper tissue (Gelsleichter 1998). The unmineralized nature of this matrix, as well as it's more open composition, would have a higher possibility of accommodating foreign infiltrative material. That the area of greater infiltration and the area of NIRS detectable age-related chemical changes are the same could likewise explain why no age-spectra correlation could be created.

As most of these proposed mechanisms for the FT-NIRS model failures were predicated on the mounting and slides, an attempt was made on some samples to explore whether the media could be sufficiently removed to allow for more useful scans. A number of organic solvents were tested for this purpose, as the precise nature of the original mounting media is unknown. Xylene and toluene failed to dissolve the resin sufficiently, but acetone was capable of removing all visual traces of the epoxy within approximately 2 hours. Scans taken of samples after the removal of the mounting media were substantially different than those taken before, with less noise and fewer peaks overall (potentially indicating a less heterogenous structure) (figure 4.5). This suggests that removal of storage media might allow old tissue sets to be analyzed with FT-NIRS in a productive manner.

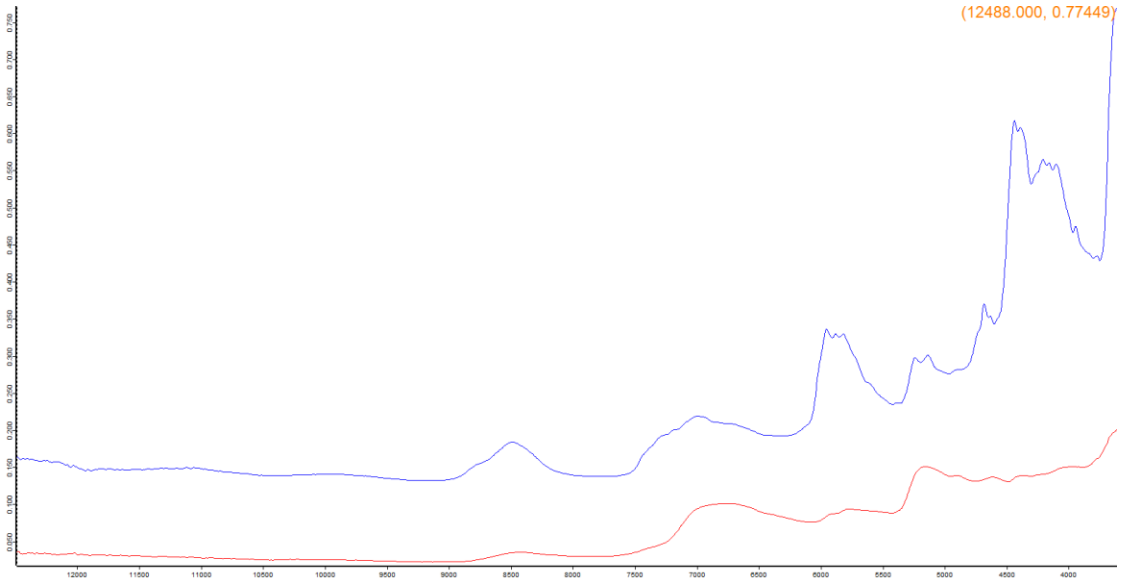


Figure 4.5. Scans taken of the same sample before (blue) and after (red) acetone cleaning, both using the 9 mm aperture.

In this case, however, a more exhaustive analysis of acetone cleaned samples was prevented. The age of the samples (20+ years) meant that, once removed from the

mounting surface, the tissue was visibly degraded and exceedingly delicate. The acetone itself also resulted in some tissue dissolution even within the time frame necessary for any noticeable impact on the resin. The end result of this was that the few samples which were successfully cleaned ($n = 4$) were either somehow broken at the time of removal from acetone ($n = 2$) or contained areas of noticeable tissue decay ($n = 2$). A more exhaustive test of the age-spectra correlations within these tissues would require a substantial portion of the original sample set to be removed from slides. Due to the age of the sample set and the high potential for irreversible sample damage, this was not performed.

Despite the lack of significant correlation, these results should not be taken as a refutation of the ability of FT-NIRS to differentiate *L. ocellata* samples by age. Instead, they strongly suggest that future FT-NIRS validation studies using skates (and, more broadly, elasmobranchs) should use a reversed order of scanning/ mounting, with scans performed before band count preparation. While this does result in a limited usefulness of already collected tissue, new studies which obtain novel sample sets will likely be strengthened by testing the conclusions here by obtaining FT-NIRS scans both before and after mounting.

WORKS CITED

Abdel-Aziz SH. 1992. The use of vertebral rings of the brown ray *Raja miraletus* (Linnaeus, 1758) off Egyptian Mediterranean coast for estimation of age and growth. Cybium 16: 121-132.

Albuquerque CQ, Lopes LCS, Jaureguizar AJ, Condini MV. 2019. The visual quality of annual growth increments in fish otoliths increases with latitude. Fisheries Research 220: 105351.

Alcalà M, Blanco M, Moyano D, Broad NW, O'Brien D, Friedrich D, Pfeifer F, Siesler H.W. 2013. Qualitative and quantitative pharmaceutical analysis with a novel hand-held miniature near infrared spectrometer. Journal of Near Infrared Spectroscopy 21: 445-457.

Andrews AH, Natanson LJ, Kerr LA, Burgess GH, Cailliet GM. 2011. Bomb radiocarbon and tag-recapture dating of sandbar shark (*Carcharhinus plumbeus*). Fishery Bulletin 109: 454-465.

Arechavala-Lopez P, Sanchez-Jerez P, Bayle-Sempere JT, Sfakianakis DG, Somarakis S. 2012. Discriminating farmed Gilthead Sea Bream *Sparus aurata* and European Sea Bass *Dicentrarchus labrax* from wild stocks through scales and otoliths. Journal of Fish Biology 80: 2159-2175.

Arrington MB, Helser TE, Benson IM, Matta ME, Gburski CM, Essington TE. 2019. Rapid age estimation of longnose skate (*Raja rhina*) vertebrae using near-infrared spectroscopy. In: Helser TE, Benson IM, Barnett BK, editors. 2019. Proceedings of the research workshop on the rapid estimation of fish age using Fourier Transform Near Infrared Spectroscopy (FT-NIRS). AFSC Processed Rep. 2019-06, 195 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Barnett BK, Benson IM, Helser TE, Passerotti M, Erickson J. 2019. Age prediction of gulf of mexico red snapper (*Lutjanus campechanus*) using near-infrared spectroscopy. In: Helser TE, Benson IM, Barnett BK, editors. 2019. Proceedings of the research workshop on the rapid estimation of fish age using Fourier Transform Near Infrared Spectroscopy (FT-NIRS). AFSC Processed Rep. 2019-06, 195 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Beamish RJ, Harvey HH. 1969. Age determination in the White Sucker. Journal of the Fisheries Research Board of Canada 26: 633-638.

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.

Beamish RJ, McFarlane GA. 1983. The forgotten requirement for age validation in fisheries biology. Transactions of the American Fisheries Society 112: 735-743.

Beamish RJ, McFarlane GA. 1985. Annulus development on the second dorsal spine of the spiny dogfish (*Squalus acanthias*) and its validity for age determination. Canadian Journal of Fisheries and Aquatic Sciences 42: 1799-1805.

Beamish RJ, McFarlane GA. 1987. History, problems, and current status: Current trends in age determination methodology. In: Summerfelt RC, Hall GA, editors. Age and growth of fish. The Iowa State University Press, Ames, Iowa. p. 15-42.

Beamish RJ, McFarlane GA. 1995. A discussion of the importance of aging errors, and an application, Walleye Pollock: the world's largest fishery. In: Secor DH, Dean JM, Campana SE, editors. Recent developments in fish otolith research. University of South Carolina Press, Columbia, SC. p. 545-565.

Beamish RJ, McFarlane GA. 2000. Reevaluation of the interpretation of annuli from otoliths of a long-lived fish, *Anoplopoma fimbria*. Fisheries Research 46: 105-111.

Begg GA, Campana SE, Fowler AJ, Suthers IM. 2005. Otolith research and application: Current directions in innovation and implementation. Marine and Freshwater Research 56: 477-483.

Bennett JT, Boehlert GW, Turekian KK. 1982. Confirmation of longevity in *Sebastes diploproa* (Pisces: Scorpaenidae) from 210Pb/226Ra measurements in otoliths. Marine Biology 71: 209-215.

Bereiter-Hahn J, Zylberberg L. 1993. Regeneration of a teleost scale. Comparative Biochemistry and Physiology A 105: 625-641.

Bestgen KR., Bundy JM. 1998. Environmental factors affect daily increment deposition and otolith growth in young Colorado Squawfish. Transactions of the American Fisheries Society 127: 105-117.

Bobelyn E, Serban A, Nicu M, Lammertyn J, Nicolai BM, Saeys W. 2010. Postharvest quality of apple predicted by NIR-spectroscopy: Study of the effect of biological variability on spectra and model performance. Postharvest Biology and Technology 55 (2010): 133-143.

Branigan PR, Meyer KA, Wahl NC, Corsi MP, Dux AM. 2019. Accuracy and precision of age estimated obtained from three calcified structures for known-age kokanee. North American Journal of Fisheries Management 39: 498-508.

Bruch RM, Campana SE, Davis-Foust SL, Hansen MJ, Janssen J. 2009. Lake Sturgeon age validation using bomb radiocarbon and known-age fish. Transactions of the American Fisheries Society 138: 361-372.

Buckmeier DL, Irwin ER, Betsill RK, Prentice JA. 2002. Validity of pectoral spines and otoliths for estimating ages of Channel Catfish. North American Journal of Fisheries Management 22: 934-942.

Buckmeier DL, Smith NG, Reeves KS. 2012. Utility of alligator gar age estimates from otoliths, pectoral fin rays, and scales. Transactions of the American Fisheries Society 141: 1510-1519.

Buckmeier DL, Sakaris PC, Schill DJ. 2017. Validation of Annual and Daily Increments in Calcified Structures and Verification of Age Estimates. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 33-79.

Burke PJ, Raoult V, Natanson LJ, Murphy TD, Peddemors V, Williamson JE. 2020. Struggling with age: Common sawsharks (*Pristiophorus cirratus*) defy age determination using a range of traditional methods. Fisheries Research 231 (2020): 105706.

Cailliet GM, Martin LK., Kusher D, Wolf P, Welden BA. 1983. Techniques for enhancing vertebral bands in age estimation of California elasmobranchs. NOAA Technical Report NMFS 8: 157-165.

Cailliet GM, Radtke RL, Welden BA. 1986. Elasmobranch age determination and verification: a review. In: Uyeno R, Taniuchi T, Matsuura K, editors. Indo-Pacific fish biology: Proceedings of the second international conference on Indo-Pacific fishes. Ichthyological Society of Japan, Tokyo. p. 345-360.

Cailliet GM, Goldman KJ. 2004. Age determination and validating in chondrichthyan fishes. In: Carrier J, Musick JA, Heithaus MR, editors. Biology of sharks and their relatives. CRC Press LLC, Boca Raton, FL. p. 399-447.

Cailliet GM, Smith WD, Mollet HF, Goldman KJ. 2006. Age and growth studies of chondrichthyan fishes: the need for consistency in terminology, verification, validation, and growth function fitting. Environmental Biology of Fishes 2006 (77): 211-228.

Cailliet GM. 2015. Perspectives on elasmobranch life-history studies: A focus on age validation and relevance to fisheries management. Journal of Fish Biology (2015) 87: 1271-1292.

Campana SE. 1984. Comparison of age determination methods for the starry flounder. Transactions of the American Fisheries Society 113: 365-369.

Campana SE, Neilson JD. 1985. Microstructure of fish otoliths. Canadian Journal of Fisheries and Aquatic Sciences 42: 1014-1032.

Campana SE, Zwanenburg KCT, Smith JN. 1990. $^{210}Pb/^{226}Ra$ determination of longevity in redfish. Canadian Journal of Fisheries and Aquatic Sciences 47: 163-165.

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.

Campana SE, Jones CM. 1998. Radiocarbon from nuclear testing applied to age validation of black drum, *Pogonias cromis*. Fishery Bulletin, U.S. 96: 185-192.

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

Campana SE. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Journal of Fish Biology (2001)59: 197-242

Campana SE, Thorrold SR. 2001. Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations? Canadian Journal of Fisheries and Aquatic Sciences 58: 30-38.

Campana SE, Natanson LJ, Myklevoll S. 2002. Bomb dating and age determination of large pelagic sharks. Canadian Journal of Fisheries and Aquatic Sciences 59: 450-455.

Campana SE, Jones C, McFarlane GA, Myklevoll S. 2006. Bomb dating and age validation using the spines of spiny dogfish (*Squalus acanthias*). Environmental Biology of Fishes (2006)77: 327-336.

Carbonara P, Bellodi A, Palmisano M, Mulas A, Porcu C, Zupa W, Donnaloia M, Carlucci R, Sion L, Follesa MC. 2020. Growth and Age Validation of the Thornback Ray (*Raja clavate* Linnaeus, 1758) in the South Adriatic Sea (Central Mediterranean). Frontiers in Marine Science 7: 586094.

Carlander KD. 1987. A history of scale age and growth studies of North American freshwater fish. In: Summerfelt RC, Hall GA, editors. Age and Growth of Fish. Iowa State University Press, Ames, Iowa. p. 3-14.

Carlson JK, Cortés E, Bethea DM. 2003. Life history and population dynamics of the finetooth shark (*Carcharhinus isodon*) in the northeastern Gulf of Mexico. Fishery bulletin 101(2): 281-292.

Casselman JM. 1987. Determination of age and growth. In: Weatherly AH, Gill HS, editors. The biology of fish growth. Academic Press, New York. p. 209-242.

Castro JI. 1993. The shark nursery of bulls bay, South Carolina, with a review of the shark nurseries of the southeastern coast of the United States. Environmental Biology of Fishes 38: 37-48.

Chang WYB. 1982. A statistical method for evaluating the reproducibility of age determination. Canadian Journal of Fisheries and Aquatic Sciences 39: 1208-1210.

Chin A, Simpfendorfer C, Tobin A, Heupel M. 2013. Validated age, growth and reproductive biology of *Carcharhinus melanopterus*, a widely distributed and exploited reef shark. Marine and Freshwater Research 64: 965-975.

Cicia AM, Driggers III WB, Ingram Jr. GW, Kneebone J, Tsang PCW, Koester DM, Sulikowski JA. 2009. Size and age estimates at sexual maturity for the little skate *Leucoraja erinacea* from the western Gulf of Maine, U.S.A. Journal of Fish Biology (2009) 75: 1648-1666.

Claiborne AM, Benson IM, Helser TE. 2019. Ageing chinook salmon (*Oncorhynchus tshawytscha*) otoliths using fourier transform near-infrared spectroscopy. In: Helser TE, Benson IM, Barnett BK, editors. Proceedings of the research workshop on the rapid estimation of fish age using Fourier Transform Near Infrared Spectroscopy (FT-NIRS). AFSC Processed Rep. 2019-06, 195 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Clement JG. 1992. Re-examination of the fine structure of endoskeletal mineralization in Chondrichthyans: implications for growth, ageing and calcium homeostasis. Marine Freshwater Research 43 (1): 157-181.

Conrath CL, Gelsleichter J, Musick JA. 2002. Age and growth of the smooth dogfish (*Mustelus canis*) in the Northwest Atlantic Ocean. Fisheries Bulletin 100: 674-682.

Copeland T, Hyatt MW, Johnson J. 2007. Comparison of methods used to age spring-summer Chinook Salmon in Idaho: validation and simulated effects on estimated age composition. North American Journal of Fisheries Management 27: 1393-1401.

Cotton CF, Grubbs RD, Daly-Engel TS, Lynch PD, Musick JA. 2011. Age, growth and reproduction of a common deep-water shark, shortspine spurdog (*Squalus cf. mitsukurii*), from Hawaiian waters. Marine and Freshwater Research 62: 811-822.

Davis CD, Cailliet GM, Ebert DA. 2007. Age and growth of the roughtail skate, *Bathyraja trachura* (Gilbert 1892). Environmental Biology of Fishes 80: 325-336.

Dean MN, Mull CG, Gorb SN, Summers AP. 2009. Ontogeny of the tessellated skeleton: insight from the skeletal growth of the Round Stingray *Urobatis helleri*. Journal of Anatomy 215: 227-239.

Dean MN, Ekstrom L, Monsonego-Ornan E, Ballantyne J, Witten PE, Riley C, Habraken W, Omelon S. 2015. Mineral homeostasis and regulation of mineralization processes in the skeletons of sharks, rays and relatives (Elasmobranchii). Seminars in Cell & Developmental Biology 46(2015): 51-67.

DiCenzo VJ, Bettoli PW. 1995. Verification of daily ring deposition in the otoliths of age-0 Spotted Bass. Transactions of the American Fisheries Society 124: 633-636.

Disspain MCF, Ulm S, Izzo C, Gillanders BM. 2016. Do fish remains provide reliable palaeoenvironmental records? An examination of the effects of cooking on the morphology and chemistry of fish otoliths, vertebrae, and scales. Journal of Archaeological Science 74: 45-59.

Driggers III WB, Hoffmayer ER. 2009. Variability in the reproductive cycle of finetooth sharks, *Carcharhinus isodon*, in the northern Gulf of Mexico. Copeia 2: 390-393.

Dwyer KS, Treble MA, Campana SE. 2016. Age and growth of Greenland Halibut (*Reinhardtius hippoglossoides*) in the Northwest Atlantic: A changing perception based on bomb radiocarbon analysis. Fisheries Research 179: 342-350.

Ebert DA, Compagno LJV. 2007. Biodiversity and systematics of skates (Chondrichthyes: Rajiformes: Rajoidei). Environmental Biology of Fishes 80: 221-237.

Elliott SAM, Carpentier A, Feunteun E, Trancart T. 2020. Distribution and life history trait models indicate vulnerability of skates. Progress in Oceanography 181: 102256.

Ellis B (ed). 1993. Chemistry and technology of epoxy resins. Springer Science+Business Media, Dordrecht, UK. 322 pages.

Feitosa LM, Dressler V, Lessa RP. 2020. Habitat use patterns and identification of essential habitat for an endangered coastal shark with vertebrae microchemistry: The case study of *Carcharhinus porosus*. Frontiers in Marine Science 125(7).

Fisher JP, Pearcy WG. 2005. Seasonal changes in growth of coho salmon (*Oncorhunchus kisutch*) off Oregon and Washington and concurrent changes in the spacing of scale circuli. Fisheries Bulletin 103(1): 34-51.

Fischer JR, Kock JD. 2017. Fin Rays and Spines. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 173-187.

Fornes V, Chaussidon J. 1978. An interpretation of the evolution with temperature of the n2+n3 combination band in water. The Journal of Chemical Physics 68: 4667-4671.

Francis MP, Maolagáin CÓ, Stevens D. 2001. Age, growth, and sexual maturity of two New Zealand endemic skates, *Dipturus nasutus* and *D. innominatus*. New Zealand Journal of Marine and Freshwater Research 35(4): 831-842.

Francis MP, Campana SE, Jones CM. 2007. Age under-estimation in New Zealand porbeagle sharks (*Lamna nasus*): Is there an upper limit to ages that can be determined from shark vertebrae? Marine and Freshwater Research 58: 10-23.

Frisk MG, Miller TJ. 2006. Age, growth, and latitudinal patterns of two Rajidae species in the northwestern Atlantic: little skate (*Leucoraja erinacea*) and winter skate (*Leucoraja ocellata*). Canadian Journal of Fisheries and Aquatic Sciences 63: 1078-1091.

Frisk MG, Shipley ON, Martinez CM, McKown KA, Zacharias JP, Dunton KJ. 2019. First Observations of Long-Distance Migration in a Large Skate Species, the Winter Skate: Implications for Population Connectivity, Ecosystem Dynamics, and Management. Marine and Coastal Fisheries 11(2): 202-212.

Fuiman LA. 2002. Special considerations of fish eggs and larvae. In: Fuiman LA, Werner RG, editors. 2002. Fishery science: the unique contribution of early life stages. Blackwell Scientific Publications, Oxford, UK.

Gallaghar MJ, Nolan CP. 1999. A novel method for the estimation of age and growth in rajids using caudal thorns. Canadian Journal of Fisheries and Aquatic Sciences 56: 1590-1599.

Geffen AJ. 1992. Validation of otolith increment deposition rate. In: Stevenson DK, Campana SE, editors. Otolith microstructure examination and analysis. Canadian Special Publication of Fisheries and Aquatic Sciences 117. p. 101- 113.

Gelsleichter JJ. 1998. Vertebral cartilage of the clearnose skate, *Raja eglanteria*: Development, structure, ageing, and hormonal regulation of growth. Dissertations, Theses, and Masters Projects. W&M ScholarWorks. Paper 1539616665.

Goldman KJ, Cailliet GM, Andrews AH, Natanson LJ. 2012. Assessing the age and growth of chondrichthyan fishes. In: Carrier JC, Musick JA, Heithaus MR, editors. Biology of Sharks and Their Relatives. Taylor and Francis Group, Boca Raton, Florida, USA. p. 423-451.

Harry AV. 2018. Evidence for systemic age underestimation in shark and ray ageing studies. Fish and Fisheries 2018(19): 185-200.

Hesler T, Benson I, Erickson J, Healy J, Kastelle C, Short JA. 2019. A transformative approach to ageing fish otoliths using Fourier transform-near infrared spectroscopy (NIRS): a case study of eastern Bering Sea walleye pollock (*Gadus chalcogrammus*). Canadian Journal of Fisheries and Aquatic Sciences 76 (5): 780-789.

Helser TE. 2019. Operationalizing FT-NIRS ageing enterprise in the national marine fisheries service: a conceptual pathway forward. In: Helser TE, Benson IM, Barnett BK, editors. Proceedings of the research workshop on the rapid estimation of fish age using Fourier Transform Near Infrared Spectroscopy (FT-NIRS). AFSC Processed Rep. 2019-06, 195 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Holden MJ, Vince MR. 1973. Age validation studies on the centra of *Raja clavata* using tetracycline. ICES Journal of Marine Science 35: 13-17.

Humason GL. 1972. Animal tissue techniques, third edition. W.H. Freeman and Company, San Fransisco. 641 pp.

Isermann DA, Meerbeek JR, Scholten GD, Willis DW. 2003. Evaluation of three different structures used for Walleye age estimation with emphasis on removal and processing times. North American Journal of Fisheries Management 23: 625-631.

James KC, Ebert DA, Natanson LJ, Cailliet GM. 2014. Age and growth characteristics of the starry skate, *Raja stellulata*, with a description of life history and habitat trends of the central California, U.S.A., skate assemblage. Environmental Biology of Fishes 97: 435-448.

James KC. 2020. Vertebral growth and band-pair deposition in sexually mature little skates *Leucoraja erinacea*: is adult band-pair deposition annual? Journal of Fish Biology 2020 (96): 4-13.

Johnson SL, Weston AJ. 1995. Temperature-sensitive mutations that cause stage-specific defects in Zebrafish fin regeneration. Genetics 141: 1583-1595.

Jolivet A, Fablet R, Bardeau JF, de Pontual H. 2013. Preparation techniques alter the mineral and organic fractions of fish otoliths: insights using Raman micro-spectrometry. Analytical and Bioanalytical Chemistry 405(14): 4787-4798.

Jones C, Brothers EB. 1987. Validation of the otolith increment aging technique for striped bass, *Morone saxatilis*, larvae reared under suboptimal feeding conditions. Fishery bulletin 86(2): 171-178.

Kalish JM. 1993. Pre- and post-bomb radiocarbon in fish otoliths. Earth and Planetary Science Letters 114: 549-554.

Kalish JM, Beamish RJ, Brothers EB, Casselman JM, Francis RICC, Mosegaard H, Panfili J, Prince ED, Thresher RE, Wilson CA, Wright PJ. 1995. Glossary for otolith studies. In: Secor DH, Dean JM, Campana SE, editors. Recent Developments in Fish Otolith Research. University of South Carolina Press, Columbia, SC. p. 723-729.

Kelly JT, Hanson JM. 2013. Maturity, size at age and predator-prey relationships of winter skate *Leucoraja ocellata* in the southern Gulf of St Lawrence: potentially an undescribed endemic facing extirpation. Journal of Fish Biology (2013) 82: 959-978.

Kemp NE. 1984. Organic matrices and mineral crystallites in vertebrate scales, teeth and skeletons. American Zoologist 24: 965-976.

Kerns JA, Lombardi-Carlson LA. 2017. History and importance of age and growth information. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 1-8.

Kerr LA, Campana SE. 2014. Chapter eleven- chemical composition of fish hard parts as a natural marker of fish stocks. In: Cadrin SX, Kerr LA, Mariani S, editors. Stock Identification Methods, 2nd edition. Academic Press, San Diego. p. 205-234.

Kimura DK, Anderl DM. 2005. Quality control of age data at the Alaska Fisheries Science Center. Marine and Freshwater Research 56: 783-789.

Kinney M, Wells R, Kohin S. 2016. Oxytetracycline age validation of an adult shortfin mako shark *Isurus oxyrinchus* after 6 years at liberty. Journal of Fish Biology 89: 1828-1833.

Klein ZB, Bonvechio TF, Bowen BR, Quist MC. 2017. Precision and accuracy of age estimates obtained from anal fin spines, dorsal fin spines, and sagittal otoliths for known-age large-mouth bass. Southeastern Naturalist 16: 225–234.

Koch JD, Quist MC. 2007. A technique for preparing fin rays and spines for age and growth analysis. North American Journal of Fisheries Management 20: 1044-1048.

Kopf RK, Davie PS. 2011. Fin-Spine Selection and Section Level Influence Potential Age Estimates of Striped Marlin (*Kajikia audax*). Copeia 2011(1): 153-160.

Kulka DW, Sulikowski J, Gedamke T. 2009. *Leucoraja ocellata*. The IUCN Red List of Threatened Species 2009.

Kusher DI, Smith SE, Cailliet GM. 1992. Validated age and growth of the leopard shark, *Triakis semifasciata*, with comments on reproduction. Environmental Biology of Fishes 35: 187-203.

Kyne PM, Simpfendorfer CA. 2010. Deepwater chondrichthyans. In: Carrier JC, Musick JA, Heithaus MR, editors. Sharks and Their Relatives II: Biodiversity, Adaptive Physiology, and Conservation. CRC Press, Boca Raton, Florida, USA. p. 37-113.

Landau S, Nitzan R, Barkai D, Dvash L. 2006. Excretal near infrared reflectance spectrometry to monitor the nutrient content of diets of grazing young ostriches (*Struthio camelus*). South African Journal of Animal Science 36(248).

Lapropoulou M, Papaconstantinou C. 2000. Comparison of otolith growth and somatic growth in two macrourid fishes. Fisheries Research 46: 177-188.

Liao H, Sharov AF, Jones CM, Nelson GA. 2013. Quantifying the Effects of Aging Bias in Atlantic Striped Bass Stock Assessment. Transactions of the American Fisheries Society (142)1: 193-207.

Long JM, Stewart DR. 2010. Verification of otolith identity used by fisheries scientists for aging channel catfish. Transactions of the American Fisheries Society 139: 1775-1779.

Long JM, Grabowski TB. 2017. Otoliths. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 189-219.

Long JM, Porta MJ. 2019. Age and growth of stocked juvenile Shoal Bass in a tailwater: Environmental variation and accuracy of daily age estimates. PLoS ONE 14(10): e0224018.

Lysaght MJ, van Zee JA, Callis JB. 1991. Laptop chemistry: a fiber-optic, field-portable, near-infrared spectrometer. Review of Scientific Instruments 62: 507.

Macnab A. 2009. Biomedical applications of near infrared spectroscopy. In: Barth A, Haris PI, editors. Biological and biomedical infrared spectroscopy. Amsterdam (Netherlands): IOS Press. p. 355-402.

Maraldo DC, MacCrimmon HR. 1979. Comparison of ageing methods and growth rates for Largemouth Bass, *Micropterus salmoides* Lacepede, from north latitudes. Environmental Biology of Fishes 4: 263-271.

Matta ME, Gunderson DR. 2007. Age, growth, maturity, and mortality of the Alaska skate, *Bathyraja parmifera*, in the eastern Bering Sea. Environmental Biology of Fishes 80: 309-323.

Maunder M, Punt AE. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research 142: 61-74.

McCathy MS, Minkley WL. 1987. Age estimation for Razorback Sucker (Pisces: Catostomidae) from Lake Mohave, Arizona and Nevada. Journal of the Arizona-Nevada Academy of Science 21: 87-97.

McEachran JD, Musick JA. 1975. Distribution and relative abundance of seven species of skates (Pisces: Rajidae) which occur between Nova Scotia and Cape Hatteras. Fishery Bulletin 73: 110-136.

McInerny MC. 2017.Scales. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 127-158.

McMillan MN, Izzo C, Junge C, Albert OT, Jung A, Gillanders BM. 2017. Analysis of vertebral chemistry to assess stock structure in a deep-sea shark, *Etmopterus spinax*. ICES Journal of Marine Science 74(3): 793-803.

McPhie RP, Campana SE. 2009. Bomb dating and age determination of skates (family Rajidae) off the eastern coast of Canada. ICES Journal of Marine Science 66: 546-560.

Meunier FJ. 2002. Skeleton. In: Panfili J, de Pontual H, Troadec H, Wright PJ, editors. Manual of Fish sclerochronology. Institut Francais de Recherche pour l'Exploitation de la Mer and Institut de Recherche pour le Developpement, Brest, France. p. 65-88.

Miller CE. 2001. Chemical Principles of Near-Infrared Technology. In: Williams P, Norris K, editors. Near-Infrared Technology in the Agricultural and Food Industries, 2nd ed. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA. p. 19-38.

Morales-Nin B. 2000. Review of the growth regulation processes of otolith daily increment formation. Fisheries Research 46(2000): 53-67.

Morison AK, Burnett J, McCurdy WJ, Moskness E. 2005. Quality issues in the use of otoliths for fish age estimation. Marine and Freshwater Research 56: 773-782.

Musick, J.A. 1999. Ecology and conservation of long-live marine animals. American Fisheries Society Symposium 23: 1-10.

Natanson LJ. 1993. Effect of temperature on band deposition in the Little Skate, *Raja erinacea*. Copeia (1993): 199-206.

Natanson LJ, Mello JJ, Campana SE. 2002. Validated age and growth of the porbeagle shark, *Lamna nasus*, in the western North Atlantic Ocean. *Fishery Bulletin* 100: 266-278.

Natanson LJ, Hamady LL, Gervelis BJ. 2016. Analysis of bomb radiocarbon data for common thresher sharks, *Alopias vulpinus*, in the north-western Atlantic Ocean with revised growth curves. Environmental Biology of Fishes 99: 39-47.

Niewinski BC, Ferreri CP. 1999. A comparison of three structures for estimating the age of Yellow Perch. North American Journal of Fisheries Management 19: 872-877.

Olson OP, Watabe N. 1980. Studies on formation and resorption of fish scales. IV. Ultrastructure of developing scales in newly hatched fry of the Sheepshead minnow, *Cyprinodon variegatus* (Atheriniformes, Cyprinodontidae). Cell and Tissue Research 211: 303-316.

Ono K, Licandeo R, Muradian ML, Cunningham CJ, Anderson SC, Hurtado-Ferro F, Johnson KF, McGilliard CR, Monnahan CC, Szuwalski CS. 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES Journal of Marine Science 72: 31-43.

Panella G. 1971. Fish otoliths: daily growth layers and periodical patterns. Science 173: 1124-1127.

Panfili J, de Pontual H, Toadec H, Wright PJ, editors. 2002. Manual of fish schlerochronology. Ifremer-IRD coedition, Brest, France. 464 pages.

Passerotti MS, Jones CM, Swanson CE, Quattro JM. 2020a. Fourier-transform near infrared spectroscopy (FT-NIRS) rapidly and non-destructively predicts daily age and growth in otoliths of juvenile red snapper *Lutjanus campechanus* (Poey, 1860). Fisheries Research 223 (2020): 105439.

Passerotti MS, Helser TE, Benson IM, Barnett BK, Ballenger JC, Bubley WJ, Reichart MJM, Quattro JM. 2020b. Age estimation of red snapper (*Lutjanus campechanus*) using FT-NIR spectroscopy: feasibility of application to production ageing for management. ICES Journal of Marine Science 77 (6): 2144–2156.

Phelps QE, Edwards KR, Willis DW. 2007. Precision of five structures for estimating age of Common Carp. North American Journal of Fisheries Management 27: 103-105.

Pierce SJ, Bennett MB. 2009. Validated annual band-pair periodicity and growth parameters of Blue-Spotted Maskray *Neotrygon kuhlii* from south-east Queensland, Australia. Journal of Fish Biology 75: 2490-2508.

Popper AN, Ramcharitar J, Campana SE. 2005. Why otoliths? Insights from inner ear physiology and fisheries biology. Marine and Freshwater Research 56: 497-504.

Portnoy DS, Hollenbeck CM, Bethea DM, Frazier BS, Gelsleichter J, Gold JR. 2016. Population structure, gene flow, and historical demography of a small coastal shark (*Carcharhinus isodon*) in US waters of the Western Atlantic Ocean. ICES Journal of Marine Science 73(9): 2322-2332.

Power G. 1978. Fish population structure in Arctic lakes. Journal of the Fisheries Research Board of Canada 35: 53-59.

Prince E, Pulos L. 1983. Proceedings of the International Workshop on Age Determination of Oceanic Pelagic Fishes: Tunas, Billfishes, and Sharks. 211 pp. NOAA Technical Report NMFS 8.

Quist MC, Jackson ZJ, Bower MR, Hubert WA. 2007. Precision of hard structures used to estimate age of riverine catostomids and cyprinids in the upper Colorado River basin. North American Journal of Fisheries Management 27: 643-649.

Quist MC, Pegg MA, DeVries DR. 2012. Age and growth. In: Zale AV, Parrish DL, Sutton TM, editors. Fisheries techniques, 3rd edition. American Fisheries Society, Bethesda, Maryland. p. 677-731.

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

Rigby CL, Wedding BB, Grauf S, Simpfendorfer CA. 2015. Novel method for shark age estimation using near infrared spectroscopy. Marine and Freshwater Research 67(5): 537-545.

Robins JB, Wedding BB, Wright C, Grauf S, Sellin M, Fowler A, Saunders T, Newman S. 2015. Revolutionising Fish Ageing: Using Near Infrared Spectroscopy to Age Fish. State of Queensland through Department of Agriculture and Fisheries.

Roggo Y, Chalus P, Maurer L, Lema-Martinez C, Edmond A, Jent N. 2007. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. Journal of Pharmaceutical and Biomedical Analysis 44(2007): 683-700.

Rude NP, Hintz WD, Norman JD, Kanczuzewksi KL, Yung AJ, Hofer KD, Whitledge GW. 2013. Using pectoral fin rays as a non-lethal aging structure for Smallmouth Bass: precision with otolith age estimates and the importance of reader experience. Journal of Freshwater Ecology 28: 199-210.

Savoy TF, Crecco VA. 1987. Daily increments on the otoliths of larval American shad and their potential use in population dynamics studies. In*:* Summerfelt RC, Hall GE, editors. The age and growth of fish. Iowa State University Press, Ames. p 413–441.

Schill DJ, Mamer ERJM, LaBar GW. 2010. Validation of scales and otoliths for estimating age of Redband Trout in high desert streams of Idaho. Environmental Biology of Fishes 89: 319-332.

Schmitt DN, Hubert WA. 1982. Comparison of cleithra and scales for age and growth analysis of Yellow Perch. The Progressive Fish-Culturalist 44: 87-88.

Schmitt PD. 1984. Marking growth increments in otoliths of larval and juvenile fish by immersion in tetracycline to examine the rate of increment formation. Fishery Bulletin, U.S. 82: 237- 242.

Schonborner AA, Boivin G, Baud CA. 1979. The mineralization process in teleost fish scales. Cell and Tissue Research 202: 203-212.

Schrank SJ, Guy CS. 2002. Age, growth, and sexual dimorphism of the Redside Dace *Clinostomus elongatus* (Kirtland) in Linesville Creek, Crawford County, Pennsylvania. Ohio Journal of Science 58: 311-316.

Secor DH, Dean JM, Laban EH. 1992. Otolith removal and preparation for microstructural analysis. In: Stevenson DK, Campana SE, editors. Otolith microstructure and analysis. Canadian Special Publication of Fisheries and Aquatic Sciences 117. p. 19-57.

Secor DH, Dean JM, Campana SE, editors. 1995. Recent developments in fish otolith research. Columbia: University of South Carolina Press.

SEDAR. 2007. SEDAR 13- Stock Assessment Report- Small Coastal Shark Complex, Atlantic Sharpnose, Blacknose, Bonnethead and Finetooth Shark. SEDAR, Silver Spring, Maryland.

Serra-Pereira B, Figueiredo I, Farias I, Moura T, Gordo LS. 2008. Description of dermal denticles from the caudal region of *Raja clavate* and their use for the estimation of age and growth. ICES Journal of Marine Science 65: 1701-1709.

Smith SE. 1984. Timing of vertebral-band deposition in tetracycline-injected leopard sharks. Transactions of the American Fisheries Society 113: 308-313.

Solberg C, Saugen E, Swenson L, Bruun L, Isaksson T. 2003. Determination of fat in live farmed Atlantic salmon using non-invasive NIR techniques. Journal of the Science of Food and Agriculture 83: 692-696.

Song Z, Zidong F, He C, Shen D, Yue B. 2009. Effects of temperature, starvation, and photoperiod on otolith increments in larval Chinese Sucker, *Myxocyprinus asiaticus*. Environmental Biology of Fishes 84: 159-171.

Song Y, Cheng F, Zhao S, Xie S. 2018. Ontogenetic development and otolith microstructure in the larval and juvenile stages of mandarin fish *Siniperca chuatsi*. Ichthyological Research 66: 57-66.

Sulikowski JA, Morin MD, Suk SH, Howell WH. 2003. Age and growth estimates of the winter skate (*Leucoraja ocellata*) in the western Gulf of Maine. Fishery Bulletin 101(2): 405-413.

Sulikowski JA, Tsang PCW, Howell WH. 2005. Age and size at sexual maturity for the winter skate, *Leucoraja ocellata*, in the western Gulf of Maine based on morphological, histological and steroid hormone analyses. Environmental Biology of Fishes (2005) 72: 429-441.

Tahvonen O, Quaas MF, Voss R. 2018. Harvesting selectivity and stochastic recruitment in economic models of age-structured fisheries. Journal of Environmental Economics and Management 92: 659-676.

Taylor GC, Weyl OLF. 2012. Otoliths versus scales: evaluating the most suitable structure for ageing Largemouth Bass, *Micropterus salmoides*, in South Africa. African Zoology 47: 358-362.

Teixeira R, Fernandéz J, Pereira J, Monteiro L. 2015. Identification of *Grapholita molesta* (Busk) (Lepidoptera: Tortricidae) biotypes using infrared spectroscopy. Neotropical Entomology 44 (129).

Urist MR. 1961. Calcium and phosphorus in the blood and skeleton of the elasmobranchii. Endocrinology 69: 778-801.

Van Dykhuizen G, Mollet HF. 1992. Growth, age estimation and feeding of captive sevengill sharks, *Notorynchus cepedianus*, at the Monterey Bay Aquarium. Australian Journal of Marine and Freshwater Research 43: 297-318.

Vance CK, Graham K, Kouba AJ, Willard ST. 2015. In vivo sex identification of the endangered Mississippi Gopher frog (*Lithobates sevosa*) using near infrared reflectance spectroscopy. Abstracts, 17th International Conference on Near Infrared Spectroscopy, Foz Do Iguassu, Brazil, OP11 (2015).

Vance CK, Tolleson DR, Kinoshita K, Rodriguez J, Foley WJ. 2016. Near infrared spectroscopy in wildlife and biodiversity. Journal of Near Infrared Spectroscopy (24): 1-25.

Vinyard EA, Frazier BS, Drymon JM, Gelsleichter JJ, Bubley WJ. 2019. Age, growth, and maturation of the Finetooth Shark, *Carcharhinus isodon*, in the Western North Atlantic Ocean. Environmental Biology of Fishes 102: 1499-1517.

Waldron ME, Kerstan M. 2001. Age validation in Horse Mackeral (*Trachurus trachurus*) otoliths. ICES Journal of Marine Science 58: 806-813.

Wedding BB, White RD, Grauf S, Tilse B, Gadek PA. 2009. Near infrared spectroscopy as a rapid, non-invasive method for sandalwood oil determination. SABRAO Journal of Breeding and Genetics 41.

Wedding BB, Forrest AJ, Wright C, Grauf S, Exley P, Poole SE. 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.

Welch TJ, can den Avyle MJ, Betsill RK, Driebe EM. 1993. Precision and relative accuracy of Striped Bass age estimates from otoliths, scales, and anal fin rays and spines. North American Journal of Fisheries Management 13: 616-620.

Whitledge GW. 2017. Morphology, composition, and growth of structures used for age estimation. In: Quist MC, Isermann DA, editors. Age and growth of fishes: principles and techniques. American Fisheries Society, Bethesda, Maryland. p. 9-31.

Wildlife and Freshwater Fisheries Division. 2009. South Carolina guide to freshwater fishes.

Williams P, Norris K, editors. 2001. Near-Infrared Technology in the Agricultural and Food Industries, 2nd ed. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA. 296 pages.

Wilson CA, Beamish RJ, Brothers EB, Carlander KD, Casselman JM, Dean JM, Jearld A, Prince ED, Wild A. 1987. Glossary. In: Summerfelt RC, Hall GE, editors. Age and Growth of Fish. Iowa State University Press, Ames, Iowa. p. 527-530.

Winkler AC, Duncan MI, Farthing MW, Potts WM. 2019. Sectioned or whole otoliths? A global review of hard structure preparation techniques used in ageing sparid fishes. Reviews in Fish Biology and Fisheries 29: 605-611.

Winton MV, Natanson LJ, Kneebone J, Cailliet GM, Ebert DA. 2014. Life history of *Bathyraja trachura* from the eastern Bering Sea, with evidence of latitudinal variation in a deep-sea skate species. Journal of the Marine Biological Association of the United Kingdom 94: 411-422.

Witt A Jr. 1961. An improved instrument to section bones for age and growth determinations of fish. The Progressive Fish-Culturist 23: 94-96.

Witten PE, Huysseune A. 2009. A comparative view on mechanisms and function of skeletal remodelling in teleost fish, with special emphasis on osteoclasts and their function. Biological Reviews 84: 315-346.

Wright PJ, Panfili J, Morales-Nin B, Geffen AJ. 2002. Otoliths. In: Panfili J, de Pontual H, Troadec H, Wright PJ, editors. Manual of Fish sclerochronology. Institut Francais de Recherche pour l'Exploitation de la Mer and Institut de Recherche pour le Developpement, Brest, France. p. 31-57.

Wright KK, Schrader W, Reinhardt L, Hernandez K, Hohman C, Copeland T. 2015. Process and methods for assigning ages to anadromous salmonids from scale samples. Idaho Department of Fish and Game, Report Number 15-03, Boise.

APPENDIX A: NIRS TABLE RESULTS

Table A.1. Results for each of the models tested (EtOH stored), including those for spectra-age correlation, spectra-length correlation, and length-age correlation. For both calibration and validation models, $\%RMSE = \left(\frac{RMSE}{Max\ age}\right) * 100$.

*For spectra-length models, $\%RMSE = \left(\frac{RMSE}{Max\ length}\right) * 100$.

Test	*n*	R^2	RMSECV	RMSEP	%RMSE	APE	Rank	Bias	RPD	Slope	Offset
All_age-spectra CV	991	92.79	3.97	-	8.02	33.04	7	0.0063	3.72	0.93	1.82
All_age-spectra .50 calibration	991	92.25	-	4.12	8.32	33.34	7	-0.0592	3.59	0.93	2.00
All_age-spectra .33 calibration	991	92.35	-	4.10	8.28	32.52	7	0.125	3.62	0.94	1.42
All_age-spectra .25 calibration	991	92.34	-	4.10	8.28	33.34	7	0.154	3.62	0.95	1.26

All_age-spectra .20 calibration	991	91.88	-	4.22	8.53	36.89	6	0.0428	3.51	0.965	0.87
All_age-spectra .10 calibration	991	91.36	-	4.35	8.79	36.33	6	0.163	3.41	0.991	0.08
Ages_1-20 Age-spectra CV	425	82.65	2.54	-	12.70	43.70	8	0.0076	2.4	0.838	1.74
Ages_1-20 Age-spectra .50 calibration	425	82.58	-	2.55	12.75	43.94	7	0.0956	2.4	0.81	1.96
Ages_1-20 Age-spectra .33 calibration	425	81.08	-	2.66	13.3	48.70	6	-0.211	2.31	0.802	2.35
Ages_1-20 Age-spectra .25 calibration	425	78.47	-	2.83	14.15	44.98	6	0.224	2.16	0.804	1.89
Ages_1-20 Age-spectra .20 calibration	425	77.49	-	2.9	14.5	59.05	7	-0.361	2.12	0.845	2.04

Ages_1-20 Age-spectra .10 calibration	425	76.01	-	2.99	14.95	58.10	5	-0.138	2.04	0.761	2.72
Ages_24-49 Age-spectra CV	566	80.46	3.28	-	6.63	6.82	7	0.0027	2.26	0.811	7.05
Ages_24-49 Age-spectra .50 calibration	566	82.88	-	3.07	6.27	6.57	7	-0.31	2.43	0.809	7.42
Ages_24-49 Age-spectra .33 calibration	566	81.07	-	3.23	6.59	6.70	7	0.104	2.3	0.815	6.79
Ages_24-49 Age-spectra .25 calibration	566	83.08	-	3.05	6.22	6.71	7	-0.521	2.47	0.819	7.27
Ages_24-49 Age-spectra .20 calibration	566	78.38	-	3.45	7.04	7.37	8	-0.182	2.15	0.846	5.93
Ages_24-49 Age-spectra .10 calibration	566	76.43	-	3.6	7.35	7.67	7	-0.542	2.08	0.834	6.73

All_age Length-Age	991	89.8	-	4.99	10.18	17.26	-	-	-	-	-
All_age Length-spectra CV	991	98.32	1.03	-	3.86*	-	7	0.0004	7.71	0.983	0.23

Table A.2. Cross validation results utilizing introduced random error in input ages. Error was randomly assigned, with an average of the listed APE and a standard deviation equal to 1/3rd APE (except in the case of APE=2). Error in ages were randomly assigned as positive or negative in the main set; (pos) indicates that the artificial error was positive (over ageing), while (neg) indicates that artificial error was negative (under ageing). APE from true is APE calculated using correct ages, while APE from simulated is APE calculated using the error-included input ages.

Simulation	R2	RMSECV	APE from true	APE from simulated	Rank	Bias	RPD	Slope	Offset
APE=1	90.27	3.6	9.58	9.72	7	-0.00123	3.21	0.905	2.989
APE=1 (pos)	91.65	3.38	9.34	9.21	7	0.00747	3.46	0.919	2.569
APE=1 (neg)	91.3	3.34	9.22	9.25	7	0.00644	3.39	0.915	2.62
APE=2	90.2	3.6	9.55	9.67	7	-0.00108	3.19	0.904	3.009
APE=2 (pos)	91.03	3.53	9.58	9.25	7	0.00765	3.34	0.913	2.784

APE=2 (neg)	91.38	3.3	9.1	9.33	7	0.00682	3.41	0.916	2.574
APE=3	89.51	3.76	9.6	10.05	7	-0.00168	3.09	0.897	3.218
APE=3 (pos)	91.24	3.52	9.92	9.46	7	0.00651	3.38	0.915	2.748
APE=3 (neg)	91.47	3.26	9.13	9.16	7	0.00711	3.42	0.917	2.517
APE=4	89.91	3.67	9.41	10.21	7	0.00893	3.15	0.902	3.072
APE=4 (pos)	92.2	3.34	9.79	8.9	8	-0.00396	3.58	0.924	2.466
APE=4 (neg)	91.23	3.28	9.17	9.31	7	0.00738	3.38	0.915	2.56
APE=5	88.87	3.9	9.38	10.35	7	0.00607	3	0.892	3.38
APE=5 (pos)	91.36	3.54	10.68	9.18	7	0.0075	3.4	0.916	2.76
APE=5 (neg)	91.82	3.14	9.01	9.02	8	0.000539	3.5	0.921	2.338
APE=6	88.55	3.98	9.34	10.53	7	0.008	2.96	0.888	3.496
APE=6 (pos)	91.18	3.62	11.12	9.24	7	0.00784	3.37	0.914	2.845
APE=6 (neg)	91.01	3.26	9.57	9.52	7	0.00662	3.34	0.912	2.573
APE=7	87.81	4.09	8.86	11.12	8	-0.000717	2.86	0.882	3.69
APE=7 (pos)	91.88	3.52	11.32	8.89	8	-0.00493	3.51	0.921	2.64

APE=7 (neg)	91.46	3.14	10.23	9.41	8	0.0071	3.42	0.918	2.383
APE=8	86.98	4.29	9.25	11.53	8	0.00641	2.77	0.875	3.919
APE=8 (pos)	92.11	3.5	11.84	8.73	8	-0.00374	3.56	0.923	2.591
APE=8 (neg)	91.91	3.02	10.17	8.84	8	-0.00164	3.52	0.922	2.258
APE=9	85.41	4.55	8.74	11.96	8	-0.00987	2.62	0.859	4.411
APE=9 (pos)	92.06	3.54	12.55	8.76	8	-0.0026	3.55	0.923	2.627
APE=9 (neg)	91.55	3.07	10.51	9.06	8	-0.00561	3.44	0.918	2.336
APE=10	83.82	4.8	8.78	12.85	8	0.00183	2.49	0.844	4.861
APE=10 rep2	83.61	4.87	10.12	13.68	8	0.00754	2.47	0.842	4.94
APE=10 rep3	84.72	4.74	8.67	12.03	8	0.00524	2.56	0.851	4.68
APE=10 rep4	84.3	4.72	8.47	12.46	8	0.00232	2.52	0.846	4.793
APE=10 rep5	84.53	4.76	8.73	13.05	8	-0.00792	2.54	0.851	4.67

APE=10 (pos)	92.12	3.57	13.26	8.82	8	-0.00554	3.56	0.924	2.64
APE=10 (neg)	91.59	3.01	11.23	9.01	8	-0.00107	3.45	0.919	2.292
APE=25	57.06	9.54	9.93	28.37	4	-0.000592	1.53	0.575	13.378
APE=25 (pos)	88.49	4.95	26.86	10.54	8	-0.00955	2.95	0.889	4.37
APE=25 (neg)	84.57	3.6	24.25	12.55	8	0.00143	2.55	0.852	3.483

Table A.3. Model results for all frozen stored sets.

Test	n	R2	RMSE	%RMSE	APE	Rank	Bias	RPD	Slope	Offset
All_age-spectra CV	1115	93.46	3.92	8	32.9	8	0.0019	3.91	0.936	1.478
All_age-spectra .50 calibration	1115	93.52	3.91	7.979592	35.31	7	-0.124	3.93	0.939	1.549

All_age-spectra .33 calibration	1115	93.57	3.89	7.938776	37.53	8	-0.117	3.95	0.93	1.747
All_age-spectra .25 calibration	1115	93.05	4.05	8.265306	39.08	8	-0.357	3.81	0.936	1.857
All_age-spectra .20 calibration	1115	93.15	4.02	8.204082	35.3	8	0.144	3.82	0.939	1.278
All_age-spectra .10 calibration	1115	92.34	4.25	8.673469	33.74	7	0.471	3.64	0.944	0.83
Ages_1-20 Age-spectra CV	577	78.65	2.62	13.1	41.18	7	0.00299	2.16	0.799	1.982
Ages_1-20 Age-spectra .50 calibration	577	80.33	2.47	12.35	40.02	7	0.000758	2.25	0.808	1.89
Ages_1-20 Age-spectra .33 calibration	577	79.15	2.55	12.75	42.52	6	0.103	2.19	0.785	2.02
Ages_1-20 Age-spectra .25 calibration	577	77.69	2.63	13.15	45.63	7	-0.262	2.13	0.852	1.719
Ages_1-20 Age-spectra .20	577	75.54	2.76	13.8	43.63	6	0.115	2.02	0.819	1.673

calibration										
Ages_1-20 Age-spectra .10 calibration	577	73.1	2.89	14.45	39.52	6	0.65	1.98	0.806	1.26
Ages_24-49 Age-spectra CV	538	87.04	2.71	5.530612	5.89	8	0.00176	2.78	0.876	4.642
Ages_24-49 Age-spectra .50 calibration	538	86.31	2.78	5.673469	5.97	8	0.25	2.71	0.829	6.16
Ages_24-49 Age-spectra .33 calibration	538	84.77	2.94	6	6.58	7	-0.122	2.56	0.881	4.58
Ages_24-49 Age-spectra .25 calibration	538	84.47	2.97	6.061224	6.49	6	0.342	2.55	0.828	6.125
Ages_24-49 Age-spectra .20 calibration	538	84.55	2.96	6.040816	6.57	7	0.00114	2.54	0.831	6.339
Ages_24-49 Age-spectra .10 calibration	538	80.03	3.36	6.857143	6.9	5	0.548	2.27	0.815	6.421
Age_12-49 Age-spectra CV	754	94.67	-	5.530612	7.65	8	0.0045	4.33	0.949	1.607

APPENDIX B: NIRS SPECTRA PRODUCED FOR EACH SAMPLE SET

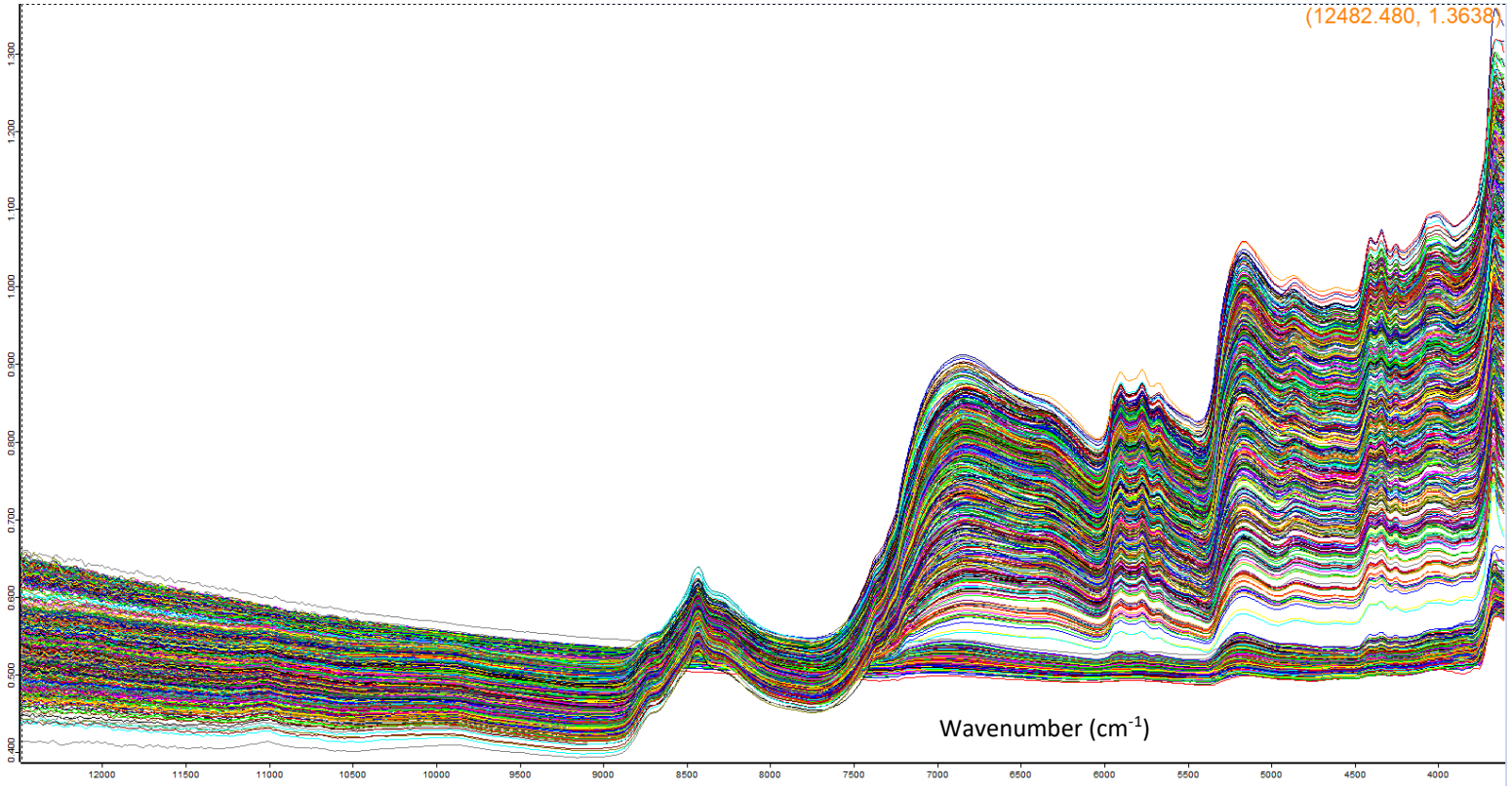


Figure B.1. Raw spectra generated from EtOH stored striped bass.

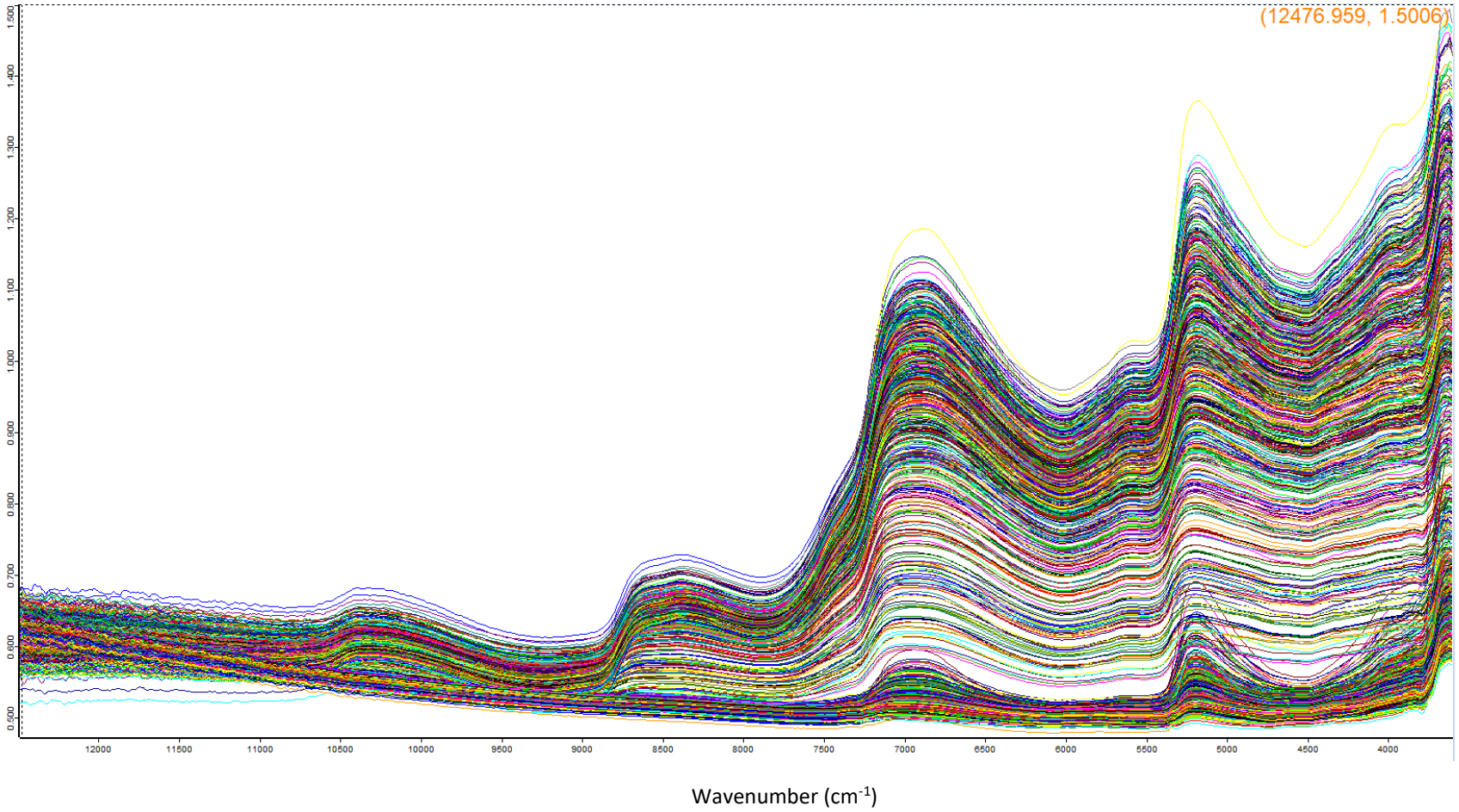
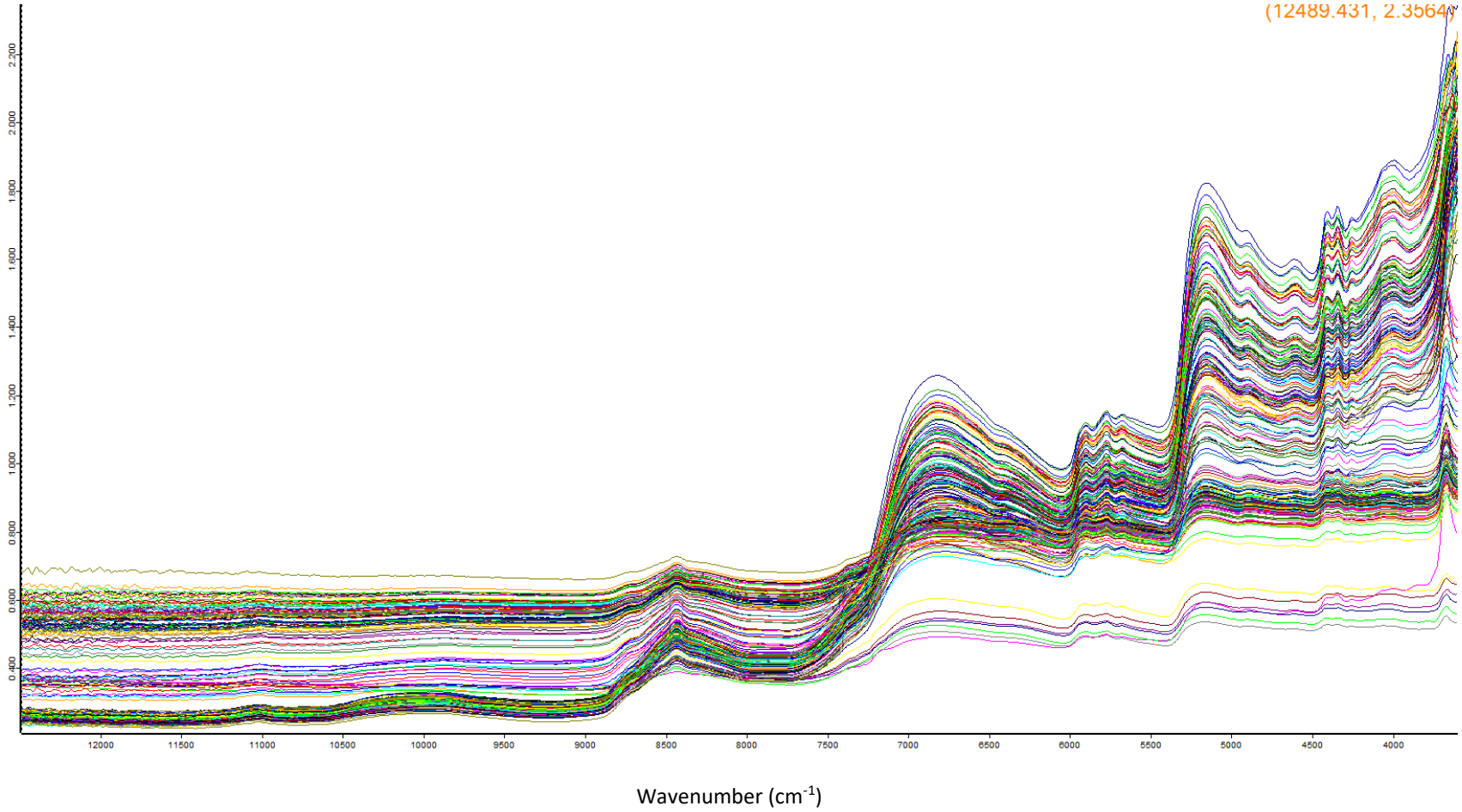
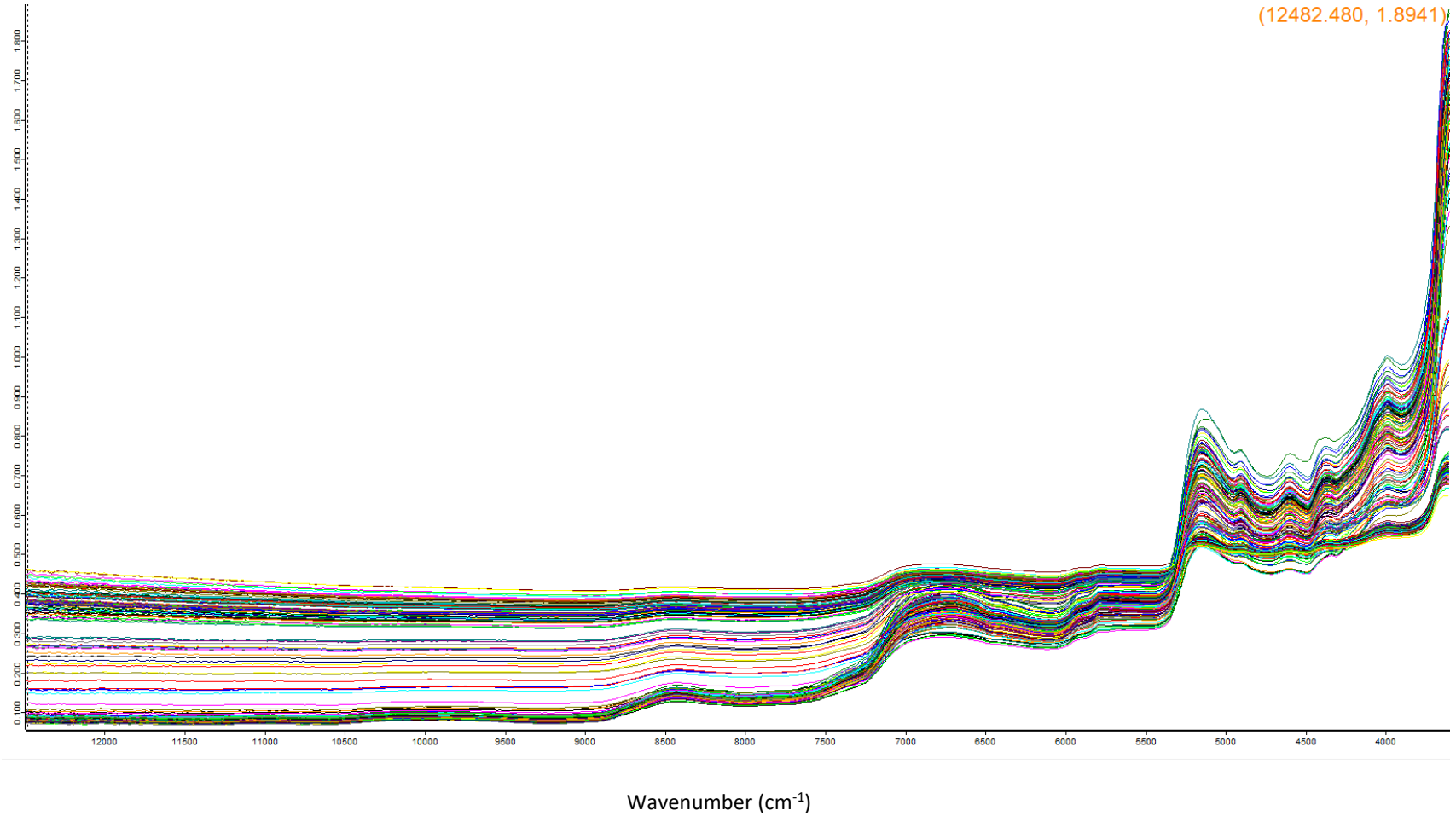


Figure B.2. Raw spectra generated from frozen stored striped bass.



B.3. Raw spectra generated from unbleached finetooth shark vertebrae.



B.4. Raw spectra generated from bleached and cleaned finetooth shark vertebrae.

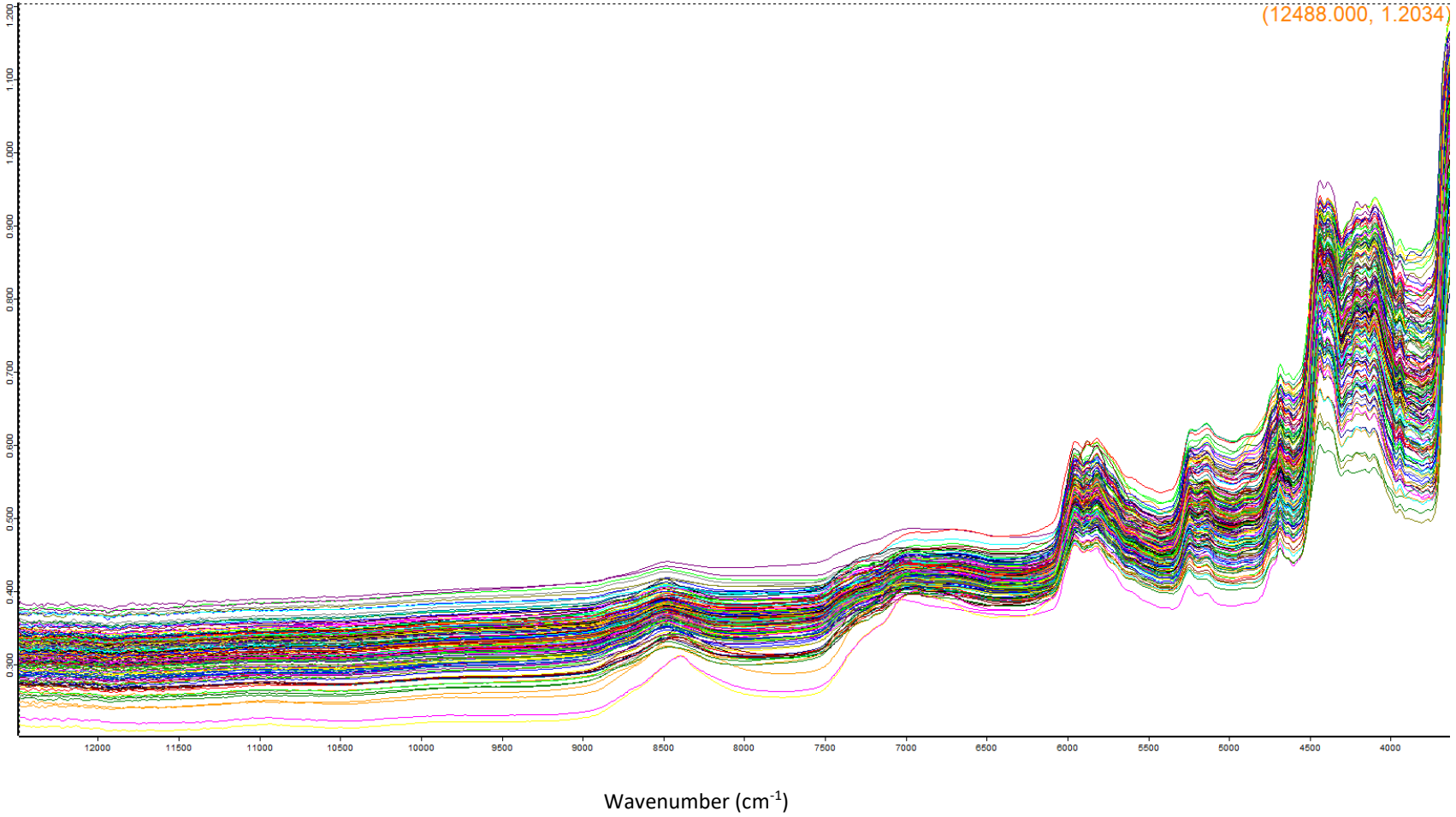


Figure B.5. Raw spectra generated from winter skate vertebrae, using a 2 mm aperture.

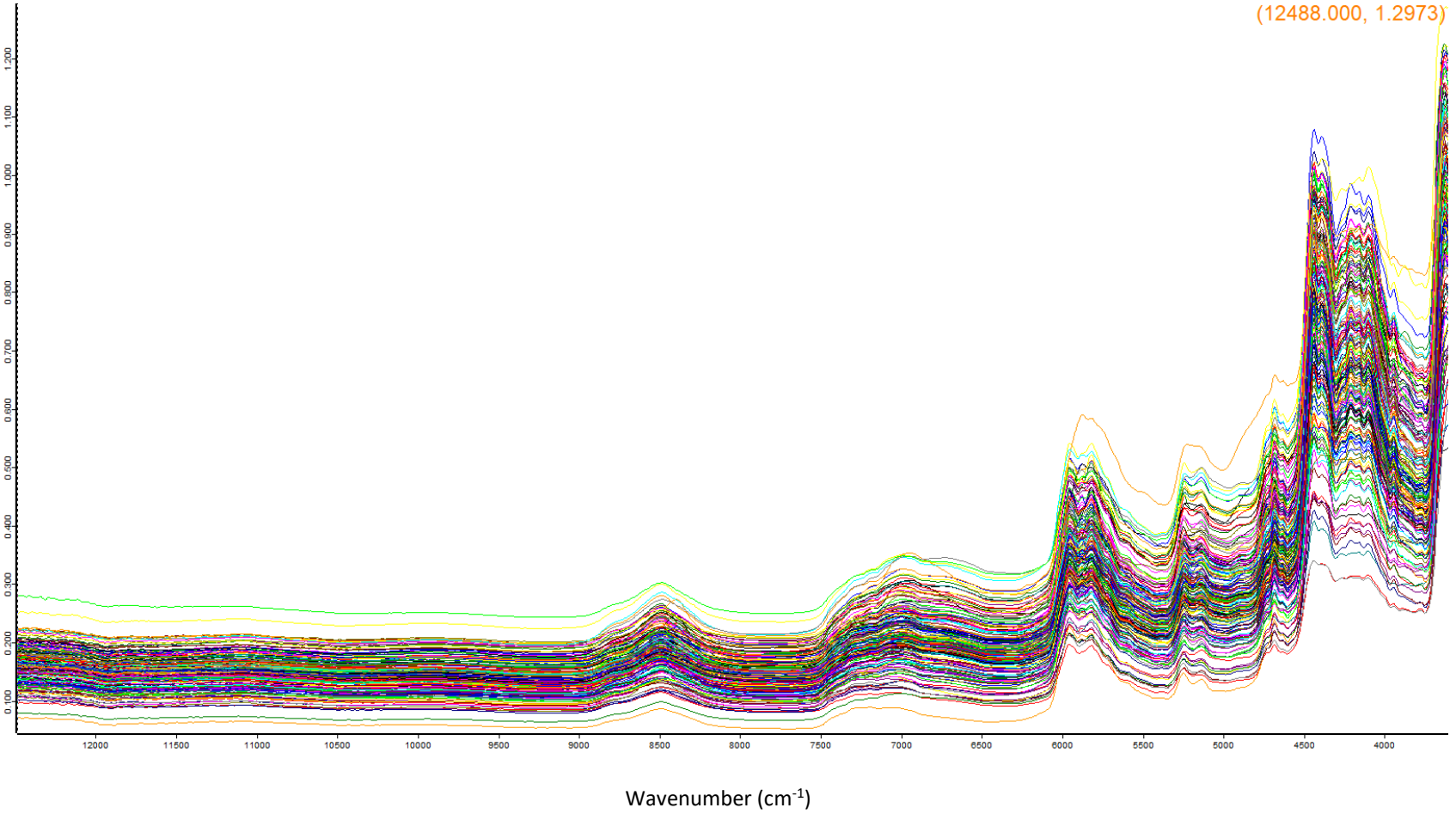


Figure B.6. Raw spectra generated from winter skate vertebrae, using a 9 mm oval aperture.