Use of Discrete Wavelet Transformations to identify groups in otolith chemistry profiles

Christian Denney*, James Hobbs, Malte Willmes, Levi Lewis
Wildlife, Fish and Conservation Biology, UC Davis. Contact: ctdenney@ucdavis.edu

Photo: Christina Parker

Introduction

- Time series analysis is becoming an increasingly widespread tool in ecological sciences. As time series become more widespread, comparing large numbers of time series will become more and more important. Making time series comparisons computationally efficient is therefore a pressing issue. In order to test efficient pattern matching and clustering techniques, we chose to use Delta Smelt otolith microchemistry profiles.
- Delta Smelt, *Hypomesus transpacificus*, is a small euryhaline fish endemic to the Sacramento-San Joaquin River Delta and is nearing extinction in the wild. Multiple factors have been identified as drivers of Delta Smelt abundance including entrainment in agriculture diversions in the South Delta. Delta Smelt are known to spawn in freshwater and rear to adulthood in brackish water resulting in two periods in the life history where individuals are vulnerable to diversion.

- To reconstruct the migration history of individuals, we use the isotopic ratios of strontium ($^{87}Sr/^{86}Sr$) deposited in the fish otolith (ear bone) to infer migration between fresh and brackish water. Otoliths are composed of calcium carbonate and protein forming daily increments. Strontium from the environment is deposited into the calcium carbonate matrix leaving a daily record of environmental history.

- Using laser ablation techniques, we can reconstruct a time series of isotope ratios to infer individual life histories. However, examining life history patterns of many individuals is challenging, time consuming, and subjective. In this study, we utilized a time series approach to efficiently characterize individual life histories.

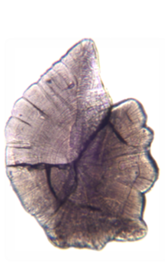
Objectives

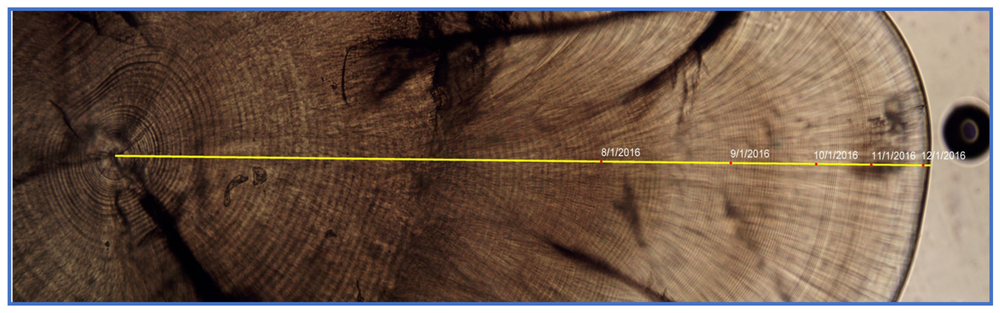
- Differentiate a large number of different time series patterns in a fast, computationally efficient manner.
- Use Delta Smelt Strontium isotope chemistry profiles to compare clustering techniques
- Clusters should represent logically consistent life history patterns
  - Freshwater Residents (FWR)
  - Brackish Water Residents (BWR)
  - Migrants (MIG)
- Differentiation of MIG based on natal origin.
- Differentiation of MIG based on timing of migration.

Methods

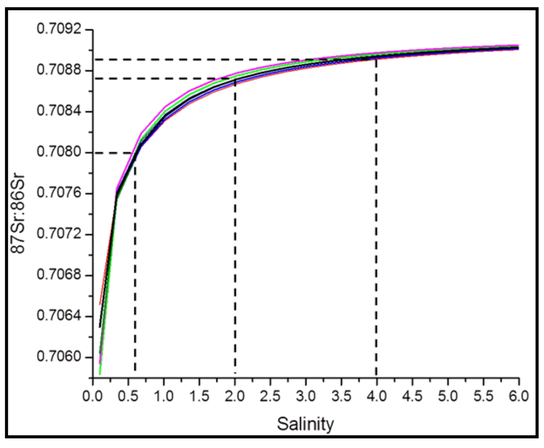
- Otoliths were analyzed according to Hobbs et al. 2010[1]
- Chemistry data was paired to increment data by matching the proportion along either the aging or the chemistry transects. (i.e. the chemistry data 50% of the distance along the transect was paired to the 90th increment on a 180 day old fish).
- Because the laser signal can degrade near the edge of the otolith and the regions of primary interest were the natal and migratory regions, only the data from the first 170 increments was used. Plots were generated using the full profiles.
- Time series were clustered using the Discrete Wavelet Transformation (DWT) technique.
- Two additional techniques were compared for calculation speed.
  - Dynamic Time Warping (DTW)
  - Symbolic Aggregate approXimation (SAX)

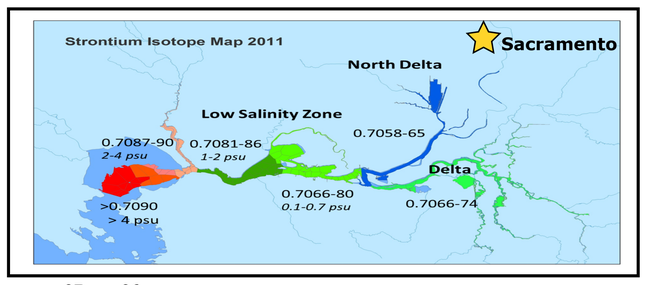
An otolith transect example





Mixing model describing the relationship between salinity and Sr isotope ratio. Each vertical line represents a sample point in the Delta. Stable ocean salinity is found at the isotopic ratio of 0.70918.



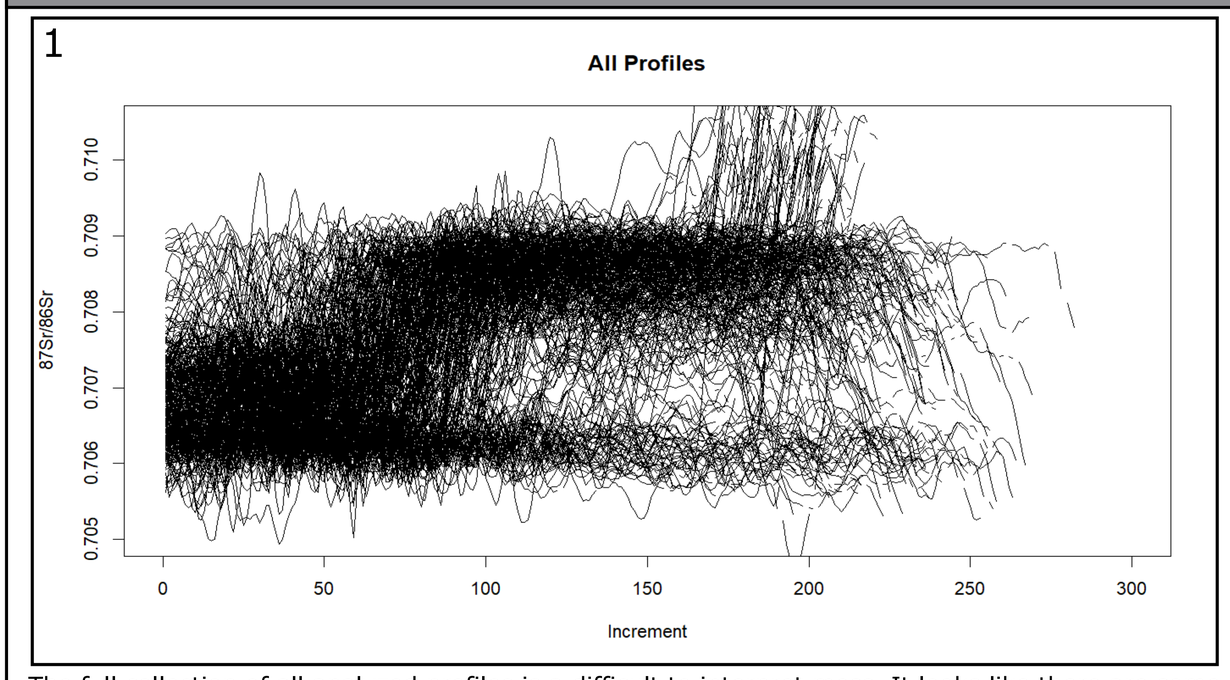


The $^{87}Sr/^{86}Sr$ ratio profile of the otolith from core to edge is used to determine migration to the low salinity zone based on the established Sr:salinity curve. Various Sr isotopic ratio values throughout the delta can be seen in this isoscape.

1: Hobbs, JA, Lewis, LS., Ikemiyagi, N., Sommer, T., Baxter., R. The use of otolith strontium isotopes ($^{87}Sr/^{86}Sr$) to identify nursery habitat for a threatened estuarine fish. (2010) *Environmental biology of fishes*. 89 (3-4), pp. 557-569.

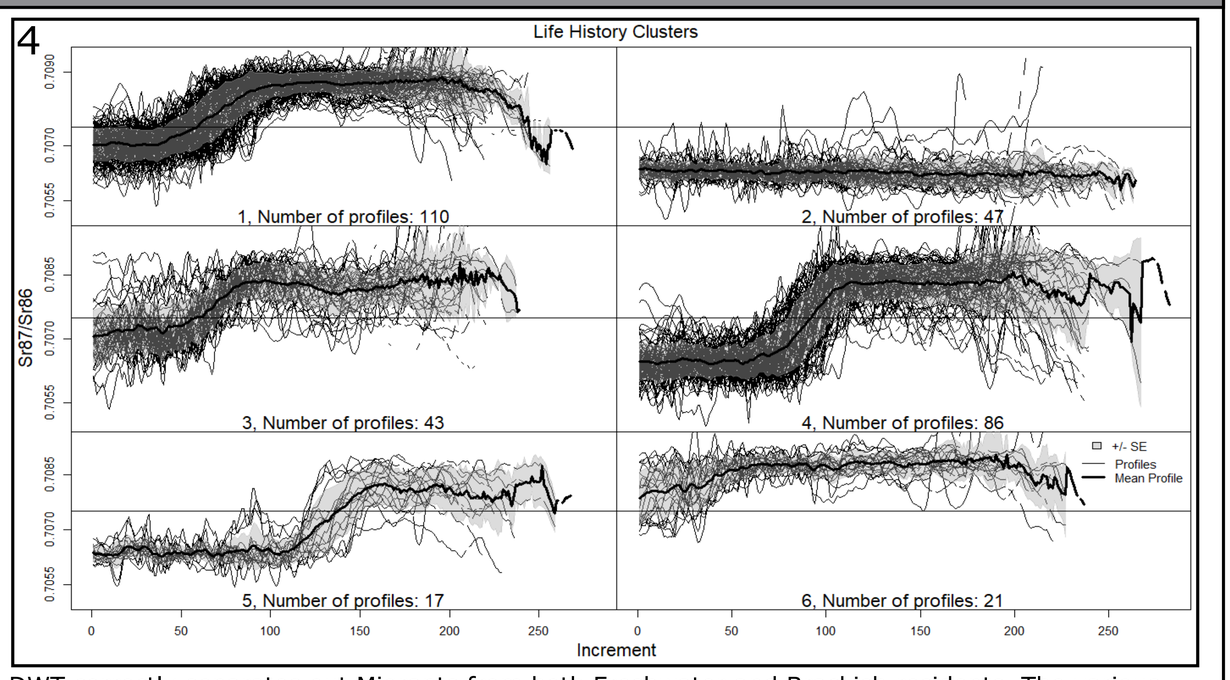
Results

1



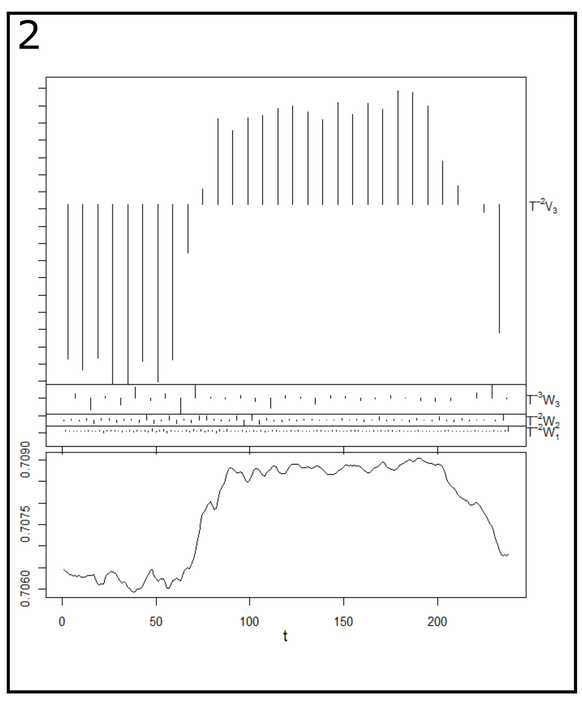
The full collection of all analyzed profiles is a difficult to interpret mess. It looks like there are some large patterns but any nuance that may exist is non-obvious. A method to quickly and efficiently collect whatever patterns may exist is necessary.

4



DWT correctly separates out Migrants from both Freshwater and Brackish residents. The various migrant groups also seem to differ in natal origin and timing of migration. Clusters are well defined.

2

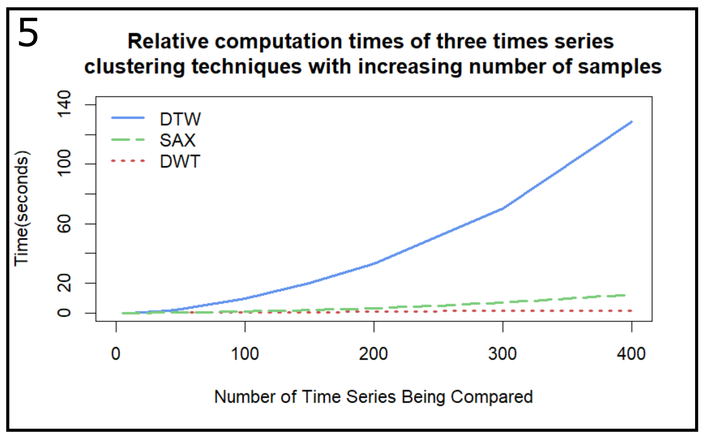


3

These are three basic life history strategies that we hope our clustering technique will find. Beyond these three basic groups, variation in natal chemistry and timing of migration would be useful group breaks

Dynamic wavelet transformation creates a series of wavelet approximations of the time series. The coefficients of these wavelet approximations are used to calculate the Euclidean distance between any two time series.

5



Both Dynamic Time Warping and Sax computation time increased exponentially with increasing number of time series while Discrete Wavelet Transformation increased linearly with comparison of increasing number of time series. In addition to being slower, the clusters created by DTW and SAX are less logically consistent than those created by DWT.

Discussion

- The timing results may not seem that substantial, even 140 seconds is not very long, but we used only 400 fish for this test while we have over 2000 in our database.
  - We collect or receive more every year
- Discrete Wavelet Transformation creates markedly better clusters while taking two orders of magnitude less time to produce those results.
  - We estimate the clustering all the Delta Smelt with DTW would take over and hour, and over 30 minutes for SAX. DWT takes less than ten seconds.
  - Clustering is usually performed many times as an exploratory technique.
- Our group is still exploring time series clustering and if you have any new techniques or suggestions on better ways to cluster, we would love your feedback.

Acknowledgments

Special thanks to Randy Baxter and Kathy Hieb at CA Dept. of Fish and Wildlife as well as to Justin Glessner at the UC Davis ICPMS Laboratory.
This project was funded by California Department of Fish and Wildlife, Agreement number P1696005