NOAA FISHERIES

Alaska Fisheries Science Center

Auke Bay Laboratories

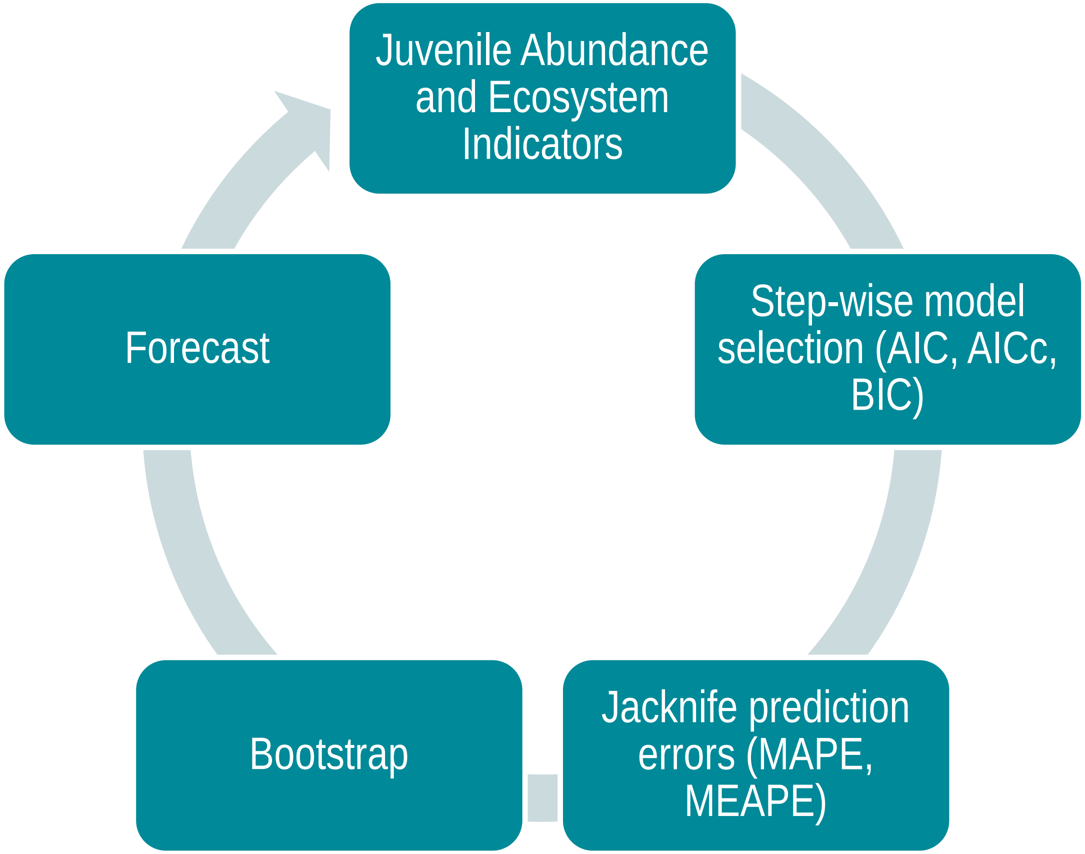
Overview of 2019 pink salmon harvest forecast model structure

Jim Murphy

Juneau, AK

Oct 19, 2018

Modelling Approach



Juvenile abundance (calibrated and trawl track distance)

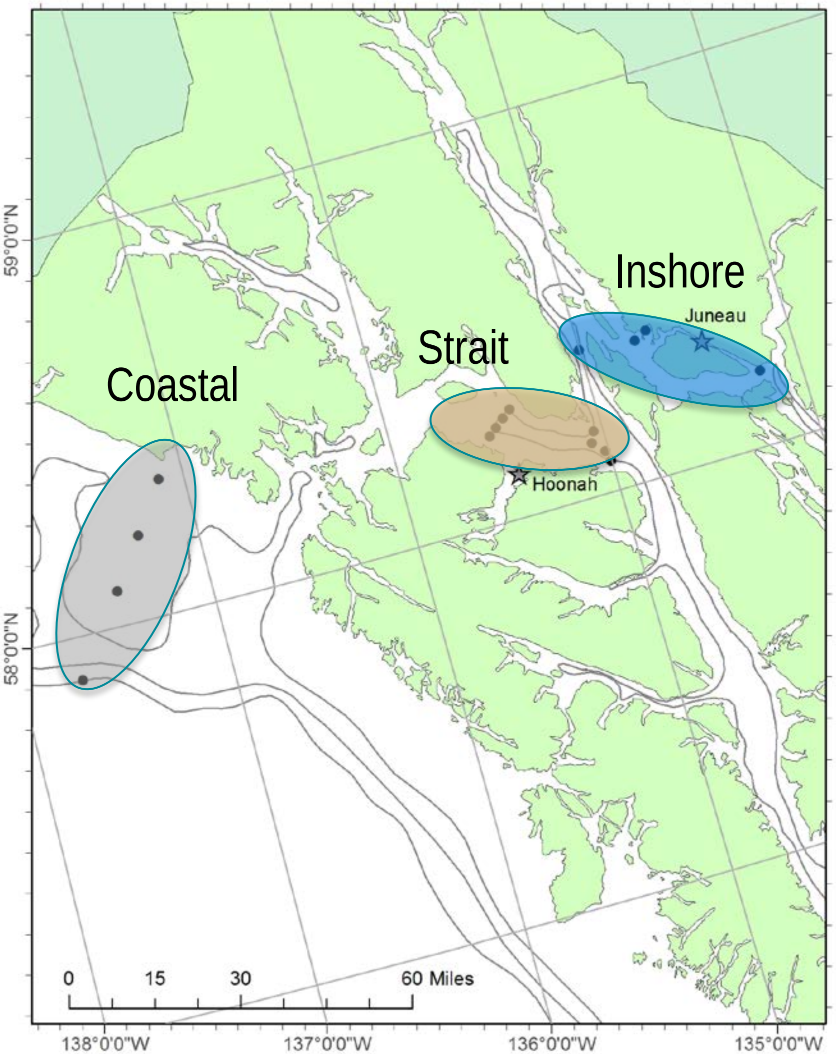
- Two approaches have been used as indices of juvenile abundance (CPUEcal and CPUEttd). CPUEttd is not considered in the 2019 forecast.
- CPUEttd: Is the catch of juvenile pink salmon divided by the distance over ground between the start and end positions of the trawl haul.
- It is complicated by tidal currents as distance over ground is not the same as distance through the water. There have been errors in the calculation of this index over time. Additional work with this index is needed before we can consider it at a viable alternative to CPUEcal.

- CPUEcal: Is the standardized catch based on a 20min trawl set. Vessel specific fishing power coefficients have been estimated from side-by-side trawling experiments.
- CPUEcal is complicated as one of the primary trawl vessels, F/V Northwest Explorer, has not been calibrated and it is assumed to have the same fishing power as the Chellissa.
- Other approaches could work, like treating vessel as a factor and modeling the vessel effect, or simply using average speed.

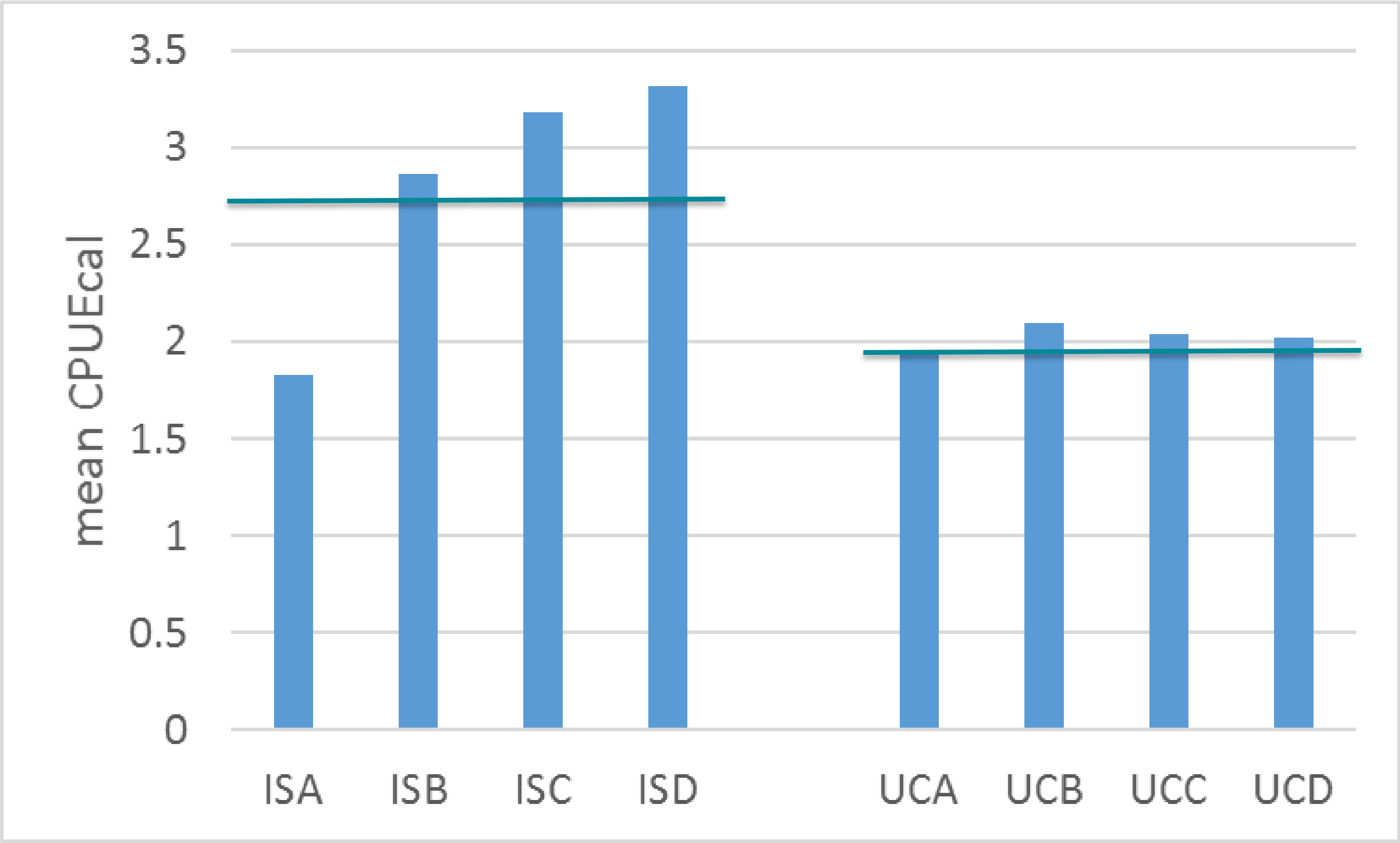
Juvenile Abundance (weighted transect means)

- Two transects are used to index the abundance of juvenile pink salmon: Icy Strait and Upper Chatham transects.
- The two transects have different means, and stations within the Icy Strait transect have different mean catches.
- Sampling effort has been inconsistent over time. If we use the average of transect means (CPUEcal_loc) rather than the overall mean of all stations, each transect is weighted equally, and standardizes sampling effort over time.
- CPUEcal_loc1 uses 3:2 weight for Icy Strait:Upper Chatham transect means, which has been a more common sampling strategy in recent years.

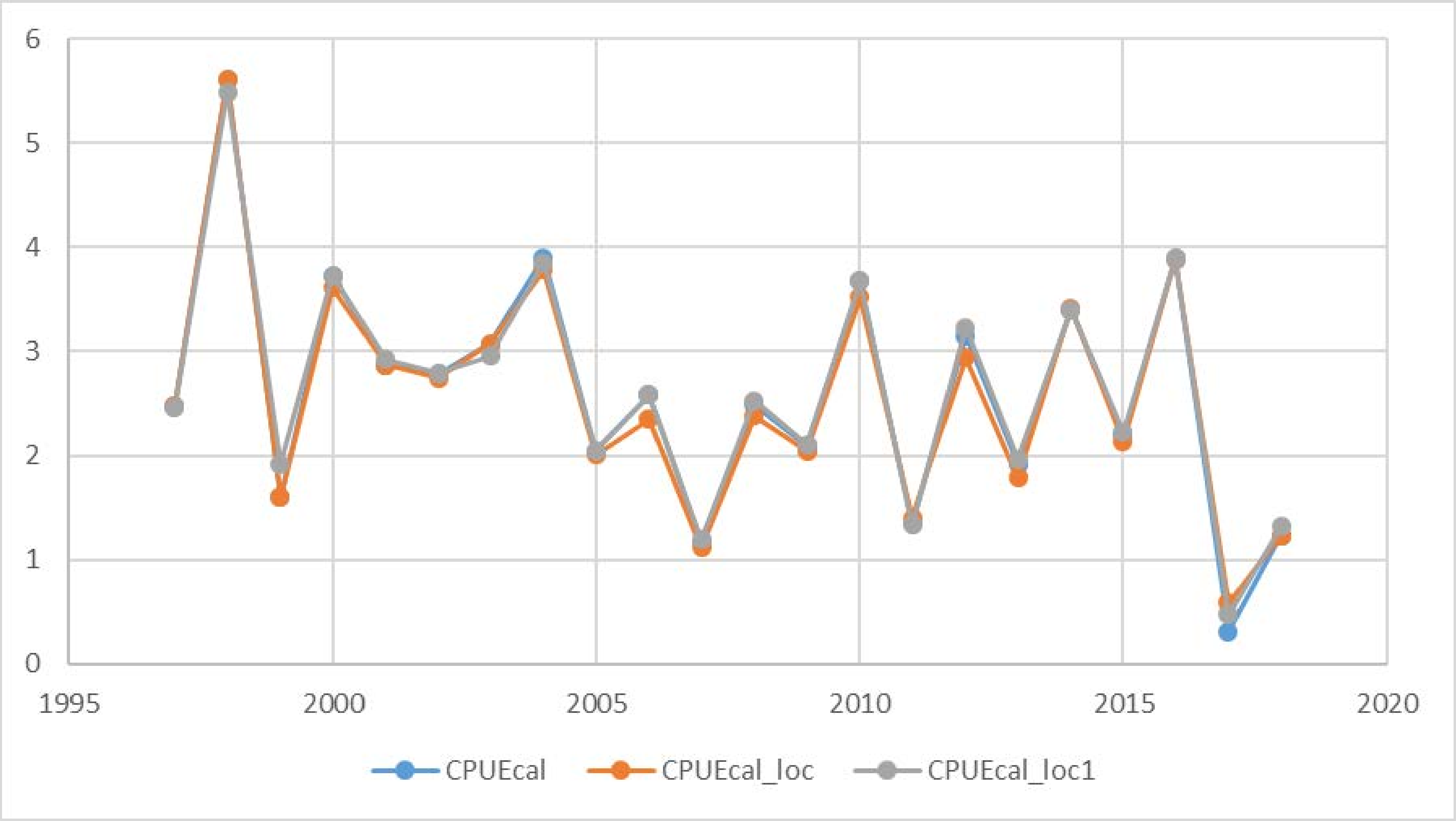
Southeast Alaska Coastal Monitoring Survey



Average ln(CPUEcal+1) by station



Unweighted and weighted transect CPUE indices



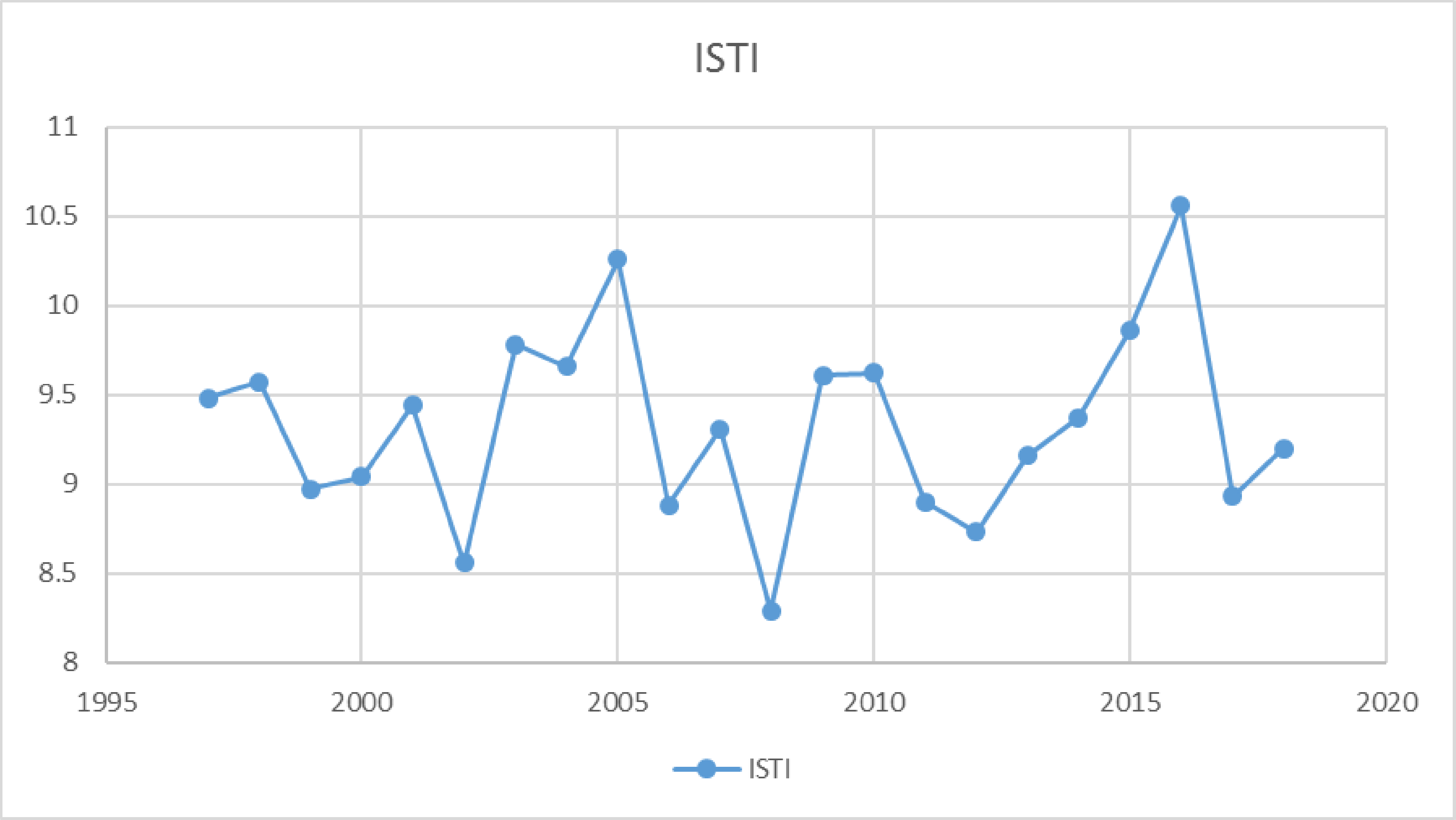
Ecosystem indicators

- A large number of indicators have been used in the past; however, a connection to pink salmon harvest forecast models has not been established.
- I recommend that we select a single forecast model and discontinue the entire step-wise model selection approach.
- We will still need a format where alternative models continue to be explored to address concerns by the forecast team.
- We may also want to consider allowing additional information other than the model output to impact the forecast.

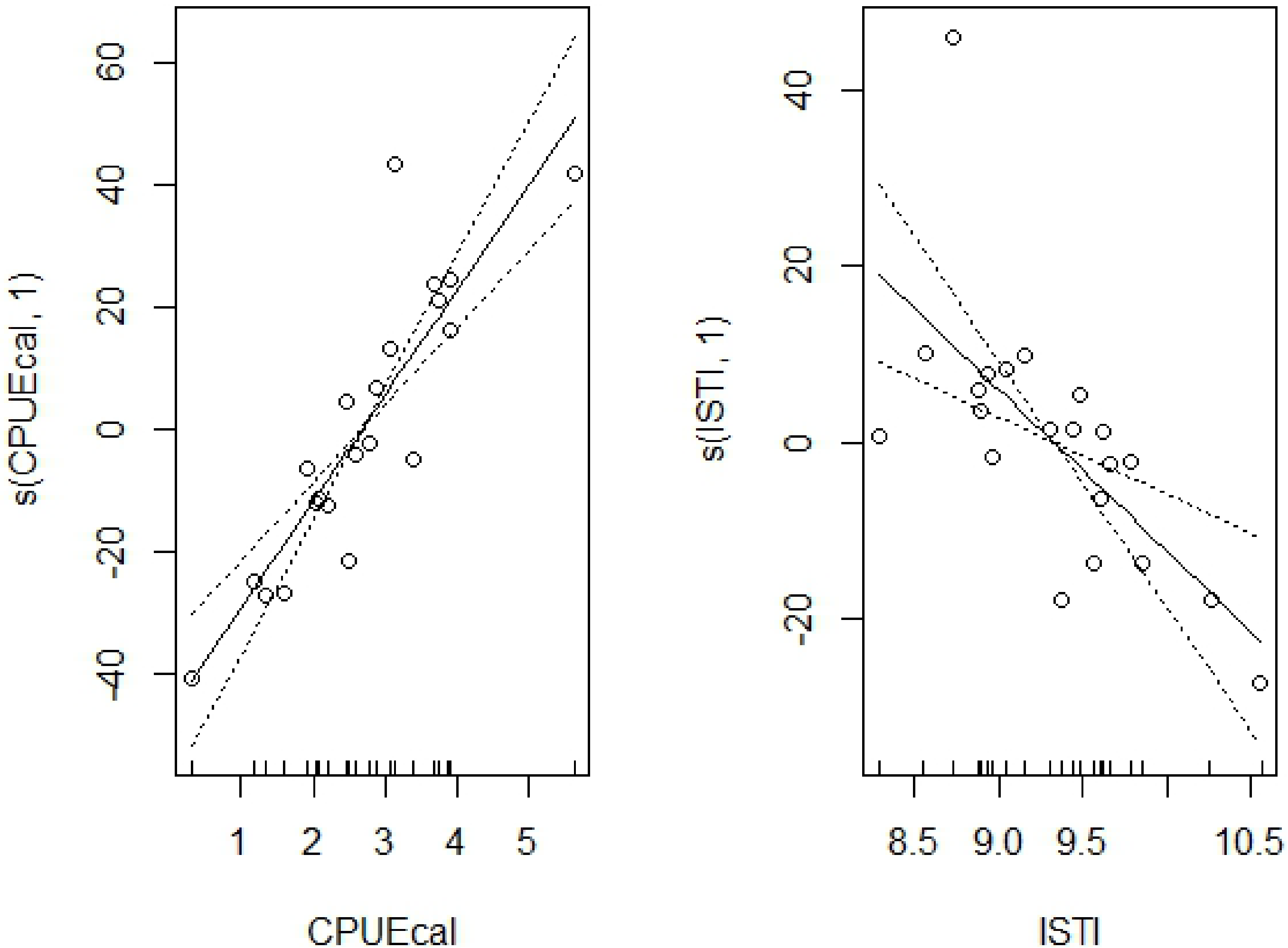
Ecosystem Indicators (ISTI)

- Average temperature in the upper 20m during May-Aug at 8 stations in Icy Strait (Icy Strait and Upper Chatham transects).
- ISTI is the one ecosystem indicator that seems to improve the fit between juvenile abundance indices and harvest.
- Why and how ISTI improves the fit is not clear. I believe it may reflect how the index is related to true abundance and not mortality.

Ecosystem Indicators (ISTI)



Partial residuals CPUEcal + ISTI



NOAA FISHERIES

Ecosystem Indicators (local)

- July24Size: estimated size on July24. Estimated from average size and date of capture from June and July SECM surveys.

- Condition: average residual of length weight regression model.

Ecosystem Indicators (Basin-scale)

- PDOsum (June, July, August) http://jisao.washington.edu/pdo/PDO.latest
- PDOwin (Nov, Dec, Jan, Feb, Mar)

- MEIwin (Nov-Mar) http://www.esrl.noaa.gov/psd/enso/mei/table.html

- ERSSTMAMJJ (GOA temperatures)
  - 55-60N, 135-150W (spring and summer, Mar-Jul)

- Temp (average of ERSST and ISTI)

Datafiles: variables

Year_Juv
SEAKCatch
NSEICatch
CPUEcal
CPUEcal_loc
CPUEcal_loc1
ISTI
ERSSTMAMJJ
MAMJJA
Temp
July24Size
Condition
MEI_win
PDO_sum
PDO_win

Datafiles: SECM2019

- # subset data by Peak month and generate list of catch by year
- cal.data<-SECM2019[SECM2019$Pink_Peak,]
- cal.data<-split(cal.data$Pink,cal.data$Year)

Step-wise model Selection (AIC,AICc, BIC)

- Initial model selection based on forward and backward stepwise model selection using AIC. Only a subset of variables are considered in the step-wise search.
- Once the model is selected based on AIC. Simpler models are evaluated for goodness-of-fit using AICc and BIC, and jackknife prediction errors (MAPE and MEAPE).
- Final model is selected based on jackknife prediction errors.

Step-wise model selection R-code

- fitcal<-step(lm(SEAKCatch~CPUEcal+ISTI+July24Size+Condition+MEI_win+PDO_sum+PDO_win, data=variables[-n,]))

Jacknife prediction error

- jacklm.reg<-function(data,model.formula,jacknife.index=0){
-   if(jacknife.index>0) {
-     var.fit<-data[-jacknife.index,]
-     var.pred<-data[jacknife.index,]
-   }
-   jack.lm<-lm(model.formula,data=var.fit)
-   predict(jack.lm,newdata=var.pred)
- }

Bootstrap

- bootstraplm<-function(cpuedata,variables,model.formula){
-     logplus1<-function(x) log(x+1)
-     catch1<-lapply(cpuedata,sample,replace=T)
-     catch2<-lapply(catch1,logplus1)
-     variables$CPUE<-unlist(lapply(catch2,mean))
-     d<-dim(variables)[1]
-     var.fit<-variables[-d,]
-     var.pred<-variables[d,]
-     boot.lm<-lm(model.formula,data=var.fit)
-     predict(boot.lm,newdata=var.pred)
- }

Boostrap summary

- boot.summary<-function(cpuedata,variables,model.formulas,model.names,quantiles=c(.1,.9)){
- boot.summary<-numeric()
- for(i in 1:length(model.formulas))
- boot.summary<-rbind(boot.summary,
- quantile(replicate(10000,
- bootstraplm(cpuedata=cpuedata, variables=variables, model.formula=model.formulas[[i]])), probs=quantiles))
- row.names(boot.summary)<-model.names
- boot.summary
- }

Forecast model summary (models)

- model.names<-c(m1='CPUE',
- m2='CPUE+ISTI',
- m3='CPUE+ISTI+Condition',
- m4='CPUE+ISTI+PDO_sum',
- m5='CPUE+ISTI+Condition+PDO_sum',
- m5='CPUE+Temp',
- m6='CPUE+Temp+Condition')
- model.formulas<-c(SEAKCatch~CPUE,
- SEAKCatch~CPUE+ISTI,
- SEAKCatch~CPUE+ISTI+Condition,
- SEAKCatch~CPUE+ISTI+PDO_sum,
- SEAKCatch~CPUE+ISTI+Condition+PDO_sum,
- SEAKCatch~CPUE+Temp,
- SEAKCatch~CPUE+Temp+Condition)

Forecast model summary function page 1

- model.summary<-function(harvest,variables,model.formulas,model.names){
- n<-dim(variables)[1]
- model.results<-numeric()
- obs<-harvest[-n]
- data<-variables[-n,]
- for(i in 1:length(model.formulas)) {
- fit<-lm(model.formulas[[i]],data=data)
- model.sum<-summary(fit)

Forecast model summary page 2

- vector.jack<-numeric()
- for(j in 1:(n-1)){
- vector.jack[j]<-jacklm.reg(data=data, model.formula=model.formulas[[i]], jacknife.index=j)
- }
- mape<-mean(abs(vector.jack-obs)/obs)
- meape<-median(abs(vector.jack-obs)/obs)

Forecast model summary function page 3

- model.pred<-unlist(predict(fit,newdata=variables[n,],se=T,interval='confidence',level=.8))
- model.results<-rbind(model.results,c(model.pred,R2=model.sum$r.squared,AdjR2=model.sum$adj.r.squared,AIC=AIC(fit),AICc=AICcmodavg::AICc(fit),BIC=BIC(fit),MAPE=mape,MEAPE=meape))
- }

Forecast model summary function page 4

- row.names(model.results)<-model.names
- dimnames(model.results)[[2]][1:3]<-c('Fit','LCI','UCI')
- model.results
- }

Forecast model output

- #generalize CPUE index
- variables$CPUE<-variables$CPUEcal

- seak.model.summary<-model.summary(harvest=variables$SEAKCatch, variables=variables, model.formulas=model.formulas, model.names=model.names)

- seak.boot.summary<-boot.summary(cpuedata=cal.data, variables=variables, model.formulas=model.formulas, model.names=model.names)

Model.summary output

	Fit	LCI	UCI	se.fit	df	residual.scale	R2	AdjR2	AIC	AICc	BIC	MAPE	MEAPE
CPUE	19.161141	12.208885	26.11340	5.236204	19	14.828562	0.5863653	0.5645951	176.7490	178.1607	179.8826	0.3195485	0.1799035
CPUE+ISTI	17.959005	12.615012	23.30300	4.016860	18	11.340246	0.7708168	0.7453520	166.3493	168.8493	170.5274	0.2000665	0.1475121
CPUE+ISTI+Condition	6.468163	-5.645229	18.58155	9.084730	17	11.048152	0.7945560	0.7583011	166.0530	170.0530	171.2756	0.2696434	0.1875443
CPUE+ISTI+PDO_sum	16.586238	11.112032	22.06044	4.105513	17	11.167650	0.7900877	0.7530444	166.5048	170.5048	171.7274	0.2765354	0.2024365
CPUE+ISTI+Condition+PDO_sum	4.705581	-7.295701	16.70686	8.977907	16	10.800055	0.8152276	0.7690344	165.8260	171.8260	172.0931	0.2646241	0.1880641
CPUE+Temp	13.566640	8.352463	18.78082	3.919282	18	10.512638	0.8030476	0.7811640	163.1665	165.6665	167.3446	0.2551138	0.1815020
CPUE+Temp+Condition	-5.084095	-16.793810	6.62562	8.781983	17	9.429849	0.8503338	0.8239221	159.4009	163.4009	164.6235	0.2130647	0.168553

Boot.summary output

	10%	90%
CPUE	16.4983524	26.200538
CPUE+ISTI	14.9020837	26.021052
CPUE+ISTI+Condition	1.0925859	18.930271
CPUE+ISTI+PDO_sum	13.4966479	24.645400
CPUE+ISTI+Condition+PDO_sum	-0.6441969	17.384270
CPUE+Temp	10.4466001	21.892260
CPUE+Temp+Condition	-10.2765854	8.144525