

Misrepresentation of Monte Carlo simulations in wildlife management: An example using piscivores (Double-crested Cormorants).

Gary W. Shugart, PhD
Slater Museum of Natural History
University of Puget Sound
Tacoma, WA 98416

Overall the FEIS is impressive in length with millions spent and an investment of countless hours by employees and volunteers for 15 years to prove that Double-crested Cormorants do indeed eat fish. In addition if salmonids are in the vicinity some will be consumed. However attempts to estimate the number of fish consumed were flawed due to a misunderstanding or misrepresentation of Methods, which have still not been detailed. Salient features of this failure include:

1. The simulation technique (Monte Carlo) used to generate consumption numbers referred to as averages or “best estimates” has been misrepresented throughout the project. The only value of the technique is to produce error bars through iterated calculations using seed values that are input to numbers generators. Seed values can be estimated, simulated, or fabricated. The values referred to as averages or “best estimates” will be the same if calculations were done once, 1,000, 10,000 or a million times. Thus, using Monte Carlo to compute average or “best estimates” of consumption was pointless and imparts no scientific or statistical validity to estimates. Basically it was an exercise in chasing your tail.
2. The purpose of the Monte Carlo simulation was to estimate error associated with the number of fish consumed based on estimated values such as population size, energy requirement, diet proportions, etc. However salmonid proportions were estimated using a previously hidden protocol involving conversion of pooled binary data to proportions. This severely limited the calculated error and was unrealistic. In addition there were numerous other caveats and built in limitations. Error was 30-50% (confidence intervals). However a partial accounting of the technique and input data appears to place these errors closer to 100-200%, which render the estimates biologically meaningless. Input data and code used for calculations, which should have been included in the FEIS, are required to assess the extent of the underestimate.
3. The FIES exposed additional problems with simulations used to concoct consumption numbers. Highlights were the manipulation of random distributions in numbers generators, admission that datasets were miniscule or simulated, and exorbitant extrapolations from few data. For example, a single detection in a stomach of the poster species for the FEIS, steelhead, was extrapolated to 1/4 to 1/3 million consumed. Since the FIES relies heavily on the bioenergetics simulation, it is critical to provide details.
4. The keystone of the FEIS remains Fredricks’ 3.x% gap in steelhead survival based in part on bioenergetics. Fredricks’ notes “there is nothing inappropriate or incorrect about either approach, however it was my opinion that the bioenergetics/smolt population approach was a better fit for the supplemental Biological Opinion”. He meant to say that there is little that is correct or appropriate in using miniscule samples, simulated data replete with numerous manipulations, and simulated consumption numbers to create an allusion of science. The guiding principle of piscivore management is that predators are inherently harmful to the resource so it really doesn’t matter what was quantified or how the quantification was done.

5. Most of the problems with the simulations in the first round of comments were ignored and are repeated below, some with elaboration.

Overview: I responded to the anonymous respondent's comments (S-2 to S-9) to some of my original comments, although most of the problems with the bioenergetics simulation were ignored or simply waved off. The crux of my analysis was that the protocol, calculations, and input dataset used for simulating and extrapolating consumption numbers were unavailable. Thus for 15 years the managers running the effort have been guided by unverified and perhaps unverifiable simulated data using undisclosed calculations and protocols. In addition it appears that few understand the basis for the simulations. However, a few details have emerged in the FEIS comments. First, the simulation used in Tillamook Bay 2012 was the same code and protocol as that used at East Sand Island. Other revelations were that data sets used for input to the simulation were very small and the putative distributions used to simulate data, were simulated or manipulated. The failure of the internal checks in calculations that I noted really aren't errors, just differences in the way numbers are computed. In the original comments, I used the Tillamook Bay 2012 simulation as a surrogate for critiquing the East Sand Island effort. Since many comments were not addressed adequately, I repeated some and expanded on others and referenced them as "Shugart 2.x".

Shugart 2.1. Monte Carlo Primer. It is clear from the citations of Furness (1978), Roby et al. (2003), and Lyons et al. (2010) and use of the "best estimated" consumption numbers throughout the FEIS that few understand that the "best estimate" consumption numbers were simulated and the simulation was superfluous for estimating "best estimates". For example from Roby et al. (2003) (bold face added)

"To obtain confidence intervals (CIs) for the output estimates from the model, a Monte Carlo simulation technique was used (Furness 1978). Each input parameter was assumed to follow a normal distribution with the measured (or assumed, where noted) mean and SE (see Results). Sampling errors in the input parameters were assumed to be uncorrelated. The Monte Carlo analysis comprised 1000 runs of the model for each year, each run using a different set of randomly generated parameter values. **In this manner, biomass and number of fish consumed were calculated for each prey type.**"

It is baffling why Monte Carlo was used to estimate biomass and number of fish since its only use is to simulate errors associated with seed values. Apparently, unrealized was that a single calculation with seed values will produce the same "best estimates" whether, 1,000, 10,000, or a million iterations. The misunderstanding with this approach is illustrated by an example and accompanying table.

Assume three variables (x, y, z) have been quantified or values could simply fabricated, where $x=10$, $y=20$, and $z=30$. Multiplied, $xyz=6000$. But with single estimate for each, or even a few estimates, it is unlikely that the phenomenon estimated by x , y & z would be adequately quantified. But ignore that requirement and plow ahead. The estimates can be used as seed values for a numbers generator and 1,000's of values can be generated. One could also just make up the iterative values, but the problem with that approach is that fabricated values tend to follow non-random patterns. So it is best to use a random numbers generator, with a distribution of the user's choice. In the example I used random normal (from Excel but also available online) to produce 1,000 values. Note that the result, $xyz=6000$, is the same "best estimate" calculated from the mean values of 1,000 iterations. A generator needs a measure of dispersion (SD or SE) and in this case I just made them up. When used appropriately with robust estimators of the seed values, the technique is valuable for estimating error in complex calculations. However whether used intentionally or unintentionally to create data or in some way to validate "best estimates", the "best

estimates” are nothing more than the single estimate calculated from a single calculation using seed values.

Thus the underlying problem with the simulated DCCO consumption numbers is that even with only single or a few observations per variable, 1,000's of values were generated giving the impression that simulations were valid. To call them “best estimates” is a clear misrepresentation and better characterizations would “only estimates”, “all we got estimates”, or “estimates resulting from a failure to understand what was being done.” With little data, Monte Carlo is nothing more than a technique to embellish data or for “What if” simulations. Referencing a Monte Carlo simulation as it is done numerous times in the FEIS should make the reader immediately ask to see the underlying seed values, SE's, and sample sizes used in the simulation. Without these data and information, simulated data are nothing more. The focus of the FEIS should include an investigation of the scientific basis for management and oversight of that effort.

Example showing that the same values are generated with a single calculation using seed values as with a 1,000 iteration Monte Carlo simulation (gray rows). Also shown is a single calculation based on error propagation that will generate similar CIs as 1,000 iterations. This illustrates that the 1,000 iterations were superfluous especially since CIs were ignored in final analysis (see Fredricks' comments and elaboration).

Single calculation	x	y	z	result (xyz)			
means (seed value)	10	20	30	6000			
SEs (seed value)	3	7	6				
						CI based on +/- 2 x SE	
				SE as proportion	SE as a value	Upper CI	Lower CI
Single calculation*							
error propagation SE, prop	0.09	0.12	0.04	0.502	3015	12030	-30
*Sqrt((3/10)^2 + (7/20)^2+(6/30)^2)							
Monte Carlo - values generated using Excel of similar						12044	-44
average 1000 iterations	10	20	30	6000			
average SE, 1000 iterations	3	7	6	3022			
Iteration 1	15	22	24	7932			
Iteration 2	11	20	34	7581			
Iteration 3	13	21	30	7932			
Iteration 4	11	21	38	8509			
Iteration 5	15	25	29	11198			
Iteration 6	10	30	34	9998			
Iteration 7	10	14	27	3894			
Iteration 8	5	35	34	6116			
Iteration 9	3	11	40	1477			
Iteration 10	10	27	26	6825			
Iteration 11	9	25	27	6411			
Iteration 12	8	22	40	7311			
Iteration 13	12	19	31	7249			
.			
.			
.			
Iteration 985	9	24	36	7430			
Iteration 986	9	27	30	7023			
Iteration 987	13	18	26	6154			
Iteration 988	10	25	20	4859			
Iteration 989	8	24	21	4005			
Iteration 990	6	28	28	4352			
Iteration 991	10	18	28	4863			
Iteration 992	8	18	31	4662			
Iteration 993	7	23	29	4931			
Iteration 994	12	21	21	5343			
Iteration 995	10	21	43	8439			
Iteration 996	8	26	35	7652			
Iteration 997	9	20	41	7686			
Iteration 998	7	26	27	4914			
Iteration 999	10	23	27	6298			
Iteration 1000	7	16	34	3766			

Other errors in the Tillamook Bay 2012 that probably occurred ESI simulations are:

Shugart 2.2. Were East Sand Island diet proportions arbitrarily pooled as was done in the Tillamook Bay 2012 simulations? I previously noted the screwy pooling of Tillamook data where 11 of 14 salmonid samples were pooled for the two month period using what was the hidden and almost mystical protocol to convert binary to quantitative diet proportions. The pooled proportions were applied to pooled Periods 1-2 & 3-4 as if independent samples, then the numbers were run as if there were four independent two-week periods. With pooling as it was done, the salmonid sample size=1 (variance is lost, but there wasn't sufficient data anyway). However, in the runs, sample sizes for salmonids were mistakenly assigned $n=14$ for pooled Periods 1 & 2, and $n=27$ for 3 & 4 (see Input Variable Values & Input Data tabs). A related mistake is using pooled sample sizes for runs of the four two-week periods. I imagine the response is, So what? DCCOs still eat salmon. True, but the concern is in science and mechanics behind the simulation or manipulation of data. With sample size dropping to 1 for salmonids, this results in the $SD=SE$ rather than $SD/\sqrt{n}=SE$ as seeds to the numbers generators, resulting in more negative values generated by the numbers generators. The bizarre admission in the FEIS comments that negative values from the number generators were simply set to zero would move the simulated "best estimates" to the right increasing proportions for salmonid take and resulting in deviation from the seed values. (One can check numbers generators by comparing the seeds to the average simulated values.)

As an example of the impact on simulations, with a proportion of .05 the "SD" is .475 using $SD="mean"*1-"mean"$ as used in the simulation. Using these to seed a random normal generator in Excel results in 466 negative values in 1,000 iterations, while a SE ($.475/\sqrt{14}$), using a sample size of 14, result in "only" 357 negatives. Apparently in the simulations, each iteration is reset to sum to 1, which violates about every concept of randomness and causes many of the proportions to deviate from seed values, but this manipulation does get rid of many negatives.

Shugart 2.3. Were adjustments made to salmonid proportions inconsistent with data? In the Tillamook Bay 2012 simulation .61 was used as the salmonid proportion in Periods 1-2 when the value should be .56 from data. ODFW personnel, no response from OSU/BRNW, indicated that the .61 provided a "better" representation of the take by cormorants. This adjustment raised the simulated take by DCCOs, and since diet proportions are relative values, it lowers the proportions for all other prey types. Since the same personnel used the same model, were similar adjustments made in East Sand Island data?

Shugart 2.4. How many stomachs were excluded from East Sand Island calculations? In the Tillamook Bay 2012, 4 of 45 stomachs were excluded because they were empty, which lacks common sense if the attempt is to sample the population. This also was a convention in the ESI sampling where it was likely done every period with variable numbers excluded (see BRNW yearly reports). An example to show the impact is illustrated by the example showing the change in proportions by including empty stomachs ($n=10$) vs excluding empty stomach ($n=8$).

	Prey 1	Prey 2	no prey
	1		
	1		
	1		
	1		
	1		
	0.5	0.5	
	0.3	0.7	
	0.7	0.3	
	0		
	0		
n=10	0.65	0.15	0.2
n=8	0.81	0.19	0

In Tillamook 2012, the excluded stomachs result in proportional change (pooled) from .252 salmonids (41 stomachs) to .229 (45 stomachs). Parsing this out to pooled Periods 1-2 & 3-4 is less dramatic, but in extrapolating out, the change is about a 10% reduction in consumption or a percent that reflects the number excluded. This mistake applies to all prey types. One could say the convention is consistent, but this is untrue unless the same proportion of stomachs were excluded from each period or conversely to the number included. This isn't a matter of interpretation, just a simple math error and a failure to think through protocol. Sampling should draw a random sample from the population not a selected sample. In parsing the biomass using prey proportions the empty stomachs should be allocated to unknown as the birds must have eaten something.

Shugart 2.5. Was Population Size variation minimized in East Sand Island simulations? On the Input Variable Values tab of the “2012TillamookBayDCCOBioenergeticsModel” Excel workbook, populations sizes used in the calculations appear in respectively labeled columns. For example, column AY, the population size for Period 1 was 21.7 pairs (indexed) reduced to 18.6 with DEE scaling ($21.6 \cdot .86$). In the simulation, the SE for each population was an artificially low value of 1.5 pairs for periods 1-3 while computed values were 17.1, 18.9, 60.8 pairs (indexed), respectively (Table 2). Period 4 only had two counts so the SD of 2.12 and resulting SE ($2.12/\text{SQRT}(2)$) was not statistically meaningful. Nonetheless, a calculated “SE” of 1.5 pairs (Table 2) was used as the seed to generate iterative values for all periods. This mistaken or intentional use of the lowest value was apparent only when the values on the Input Variable Values tab (i.e., raw input data) were examined. Note that if the actual SEs were used, this would generate more negative values for lower tails of some CIs. Thus for purposes of the simulation, the SEs used to generate the CIs, were simulated and artificially low for Periods 1-3 and not statistically valid for Period 4. Was this done purposefully in order to minimize CIs and was similar rule used in the code used for East Sand Island calculations?

Shugart 2.6. At what point did genetic analysis enter into the calculations at East Sand Island? When was the hidden protocol for converting binary data to decimal proportions adopted? How did this affect the results?

Shugart 2.7. Were Daily Energy Expenditure (DEE) overestimated and variation underestimated?

DEE estimates were from Lyons et al. 2010, Table 3.1 based on a grand mean of seven males and three females from Rice and East Sand Islands over five years using non-standard day lengths (i.e., the inclusion of extra day light hours). To what extent did the inclusion of extra daylight hours increase DEE and resulting projection of total biomass? In addition in the Tillamook Bay 2012 simulation, variation used to seed numbers generators was 136.5 kJ which was artificially low and should be 275 kJ using the SD's computed from Lyons et al 2010 using a standard formula to combine variances. Again was this value underestimation just another error?

Shugart 2.8. Were estimates for energy density and prey masses, both critical in extrapolations to number of prey, simulated (see original comments). Values used for calculations to convert biomass for each prey type to grams of prey to number of prey type appear to be simulated. In both matrices for Tillamook Bay 2012, 11 of 20 values were based on sample size of 1. For these, SD is artificially set at 20% of the value. Is this correct?

Shugart 2.9. What were raw input data to East Sand Island simulations by period? These data need to be provided by period since total consumption was summed subtotals of consumption by period. I've pointed to the problem of minimal sample size as a result of parsing data into too many categories and periods in previous comments. A fundamental problem with the stomach analysis approach is illustrated by the Tillamook Bay 2012 data. The seed values, termed means, that are fed into the simulation were based on minimal data sets, in six instances "means" were based on a single value (see table). For these, the SEs used to seed the numbers generator was simulated. In addition to minimal data, the seeds are assumed to be normal estimates of the values. However, this is clearly an invalid assumption given the number of zeros in the data. For example a value estimated with a single item in one stomach, was used as a mean for a data array with 14 zeros once in Period 1-2 and 22 zeros and one value five times in Periods 3-4. Clearly these are non-normal data but inexplicably were assumed to be normal for purposes of the simulations. The relevant point is that this underscores a lack of data and invalid assumptions were of no concern to the simulators. Based on the greater number of periods and prey types in East Sand Island simulations, the situation there probably was worse.

Table summarizing data used as input for Tillamook Bay 2012 simulations. Salmonid proportions were estimated separately using a pooled approach resulting in an additional five samples sizes of 1 for each pooled period.

Number of values used to compute seed values termed means and SEs	Period 1-2 (n=15)	Period 3-4 (n=30)
1	1	5
2	1	4
3	2	3
4	1	
5		
6		1*
7		
8	1*	
9		
10		
11		
12		1
13		
14		1
Total values estimated	6	15

*overall salmonids parsed into five categories in the simulations

GW Shugart’s responses to anonymous review comments in FEIS, labeled S-2 to S-9. A fundamental problem with the stomach analysis approach is illustrated by the Tillamook Bay 2012 data. I

S-2 Comment. Errors were found in how variance estimates were calculated using the bioenergetics model and approach. These variance estimates were used in NOAA Fisheries’ “survival gap” analysis. Use of these estimates greatly increases uncertainty surrounding DCCO consumption estimates and would drastically change the RPA action 46 target colony size identified as the management objective.

S-2 Anonymous respondent. Bioenergetics Model and DCCO Salmonid Consumption Estimates S-2 Errors were found in how variance estimates were calculated using the bioenergetics model and approach. These variance estimates were used in NOAA Fisheries’ “survival gap” analysis. Use of these estimates greatly increases uncertainty surrounding DCCO consumption estimates and would drastically change the RPA action 46 target colony size identified as the management objective. Variance measures for bioenergetics based consumption estimates have been updated in the FEIS (see S-3). NOAA Fisheries’ “survival gap” analysis used the best (here, median) per capita consumption estimates. When the error was corrected and the model rerun, the point estimates did not change substantially (resulting in a 3.54% gap in steelhead survival as presented in the FEIS Section 1.2, compared to the 3.65% as presented in the DEIS). The slight change in the gap estimate was largely due to the model simulation process.

Bioenergetic consumption estimates updated with revised variance calculations in Chapter 1 and Appendix C. FEIS Figures 1-3 and 1-7 revised and include updated confidence intervals.

S-3 Anonymous respondent. The diet proportion standard deviation formula used in the bioenergetics model for Tillamook Bay analysis was incorrect. This error appears to be inherent to the model and was likely applied to East Sand Island DCCO consumption data, resulting in underestimation of the variation and resulting See S-2. This error was also present in the calculations performed for the DEIS. The formula has been corrected and revised consumption estimates and 95% confidence intervals calculated using the corrected uncertainty estimates and are presented in the FEIS (Appendix C). This error does not have a direct effect on the best estimates of smolt consumption, but the revised confidence intervals in the FEIS are wider

G. W. Shugart additional comment: Since I didn't mention variance I assume this was misdirected or fails in attempting to paraphrase something. Nonetheless, variance was not considered in the model or consumption numbers except in an offhand manner in the context of SDs and SEs, which are simulated/fabricated in many instances, especially for proportional consumptions, which appear to be based on miniscule samples (see Appendix C, Table 5, Fig. 5 below). Note the comment Appendix C added that Monte Carlo can be run "irrespective of sample size" technically is true. In fact the model can be run with single simulated values for input or by simply fabricating data, making it seem as if the simulation was based in reality. There is simply no confidence in the output of the model and until the raw input values are provided, consumption numbers remain as "garbage in, garbage out" simulation.

I included a data table from the Tillamook Bay 2012 which we now know based on responses in the FEIS followed the same protocol used at East Sand Island. The respondent doesn't appear to have a clue about how the original consumption numbers were generated. The 3.6% gap was based on the CI midpoints, designated "best estimates". Again they appear best because they peak as result of misapplication of normal distribution (see below). Given few data, a simple random distribution should have been used. Or more simply, just pick some random numbers between .0001-.05 (proportions, not percents) to fill in the seed values for the numbers generators. After all, everyone knows DCCOs eat salmonids so the goal is to cook up rationale for control. Please show the data and the code used for calculations.

The following data table is from the input for Tillamook Bay 2012 simulation of DCCO consumption. Salmonid vs non-salmonid categories are shown in the header. Note that the means consisting mostly of zeros were used as estimators of normal means. In six cases means were based on single values. The salmonid proportions pooled for Periods 1-2 & Periods 3-4 were parsed after frequency of occurrence analysis then conversion of binary data to proportions precise to three decimals. The same type of simulation was used at East Sand Island but the data for individual periods is unavailable. In summary of the data table, those running the model for DCCOs don't have an understanding or ignore tenets of statistical sampling and estimating biological phenomenon.

	stomach	ID'able gms	Salmonid Mass	Clupeid Mass	Anchovy Mass	Osmerid Mass	Surfperch Mass	Pleuronectid Mass	Stickleback Mass	Sculpin Mass	Sandlance Mass	Gunnel Mass	Snake Prickleback Mass	Prickleback Mass	Prickleback/Gunnel Mass	Pipefish Mass	Rockfish Mass	Greenling Mass	Lingcod Mass	Unid Nonsal Mass	Crust
Period 1	1	82.6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	48.9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Period 2	3	206.1	0	0	0	0	0	0	0	0.120	0	0	0	0.586	0	0	0	0	0	0.201	0.093
	4	387.8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	5	19.5	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	6	46.5	0	0	0	0	0	0	0	0.989	0	0	0	0	0	0	0	0	0	0	0.011
	7	127.4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	9	27.7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	10	51.5	0.856	0	0	0	0	0	0	0.012	0	0	0	0	0	0.072	0	0	0	0.060	0
	11	169.6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	12	186.5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	13	excluded																			
	14	25.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	15	76.7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pooled 1-2 proportions			0.561	0	0	0	0	0	0	0.152	0	0	0	0.113	0.000	0.005	0	0	0	0.090	0.079
Pooled 1-2 count		8							4				2		1					3	3
Period 3	16	4.9	0	0	0	0	0	0	0	0	0	0	0	0	1.000	0	0	0	0	0	0
	17	35.4	0.893	0	0	0	0.107	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	18	202.3	0	0	0	0	0	0	0.037	0.000	0	0	0	0	0	0.857	0	0	0	0.107	0
	19	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	20	113.9	0	0	0	0	0	0	0	0	0	0.225	0	0	0	0	0	0.509	0	0.266	0
	21	22.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	22	113.5	0.058	0.907	0	0	0.029	0	0	0	0	0	0	0	0	0	0	0	0	0.005	0
	23	26.7	0.011	0	0	0	0.768	0	0	0	0	0	0	0.101	0.000	0	0	0	0	0.120	0
	24	11	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	25	25.8	0	0	0	0	0.570	0	0	0	0	0	0	0	0	0	0	0	0	0.283	0.147
	26	17	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	27	136.3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	28	12.5	0	0	0	0.608	0	0	0	0	0	0	0	0	0	0	0	0	0	0.392	0
	29	117.3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	30	63.3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
	31	excluded																			
	32	excluded																			
	33	237.4	0	0	0	0	0	0	0	0	0	0	0.816	0	0	0	0	0	0	0.184	0
	34	101.3	0	0	0.657	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.343	0
	35	372.2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Period 4	36	142.1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000	0
	37	44	0	0	0	0.605	0	0	0	0	0	0	0	0	0	0	0	0	0	0.395455	0
	38	57.1	0.196	0.266	0	0	0.525	0	0	0	0	0.012	0	0	0	0	0	0	0	0	0
	39	excluded																			
	40	52.1	0	0	0	0	0	0.050	0.950	0	0	0	0	0	0	0	0	0	0	0	0
	41	33.8	0	0	0	0.959	0	0.041	0	0	0	0	0	0	0	0	0	0	0	0	0
	42	82.2	0	0	0.106	0.557	0	0.018	0	0	0	0	0	0	0	0	0	0	0	0.319	0
	43	22.8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	44	11.5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	45	68.9	0.308	0	0	0	0.147	0	0	0	0	0	0.234	0	0	0	0	0	0	0.312	0
Pooled 3-4 proportions			0.091	0.043	0.028	0.037	0.329	0	0.004	0.074	0	0.008	0.009	0.104	0.041	0	0.032	0.056	0	0.138	0.005
Pooled 3-4 count		6	2	2	1	14		3	3		1	2	3	1		1	2			12	1

S-3. G. W. Shugart additional comment: Nothing was done as far as I can determine.

S-4. G. W. Shugart additional comment: I know what Monte Carlo does (see Monte Carlo Primer above). One problem with the approach, and I suspect it the case for all DCCO data, is that the simulation is being used to “launder” sparse data making it seem as if there were robust estimators for input variables. At least this is the case for Tillamook Bay 2012. If data are lacking for a single calculation, doing it 1,000 times does not improve data. The respondent doesn’t appear to realize that Monte Carlo can be run with no or little data. Just detect a few fish in a few stomachs, use these to estimate diet proportions expressed as means and SEs which are used as seed values to run the calculations 1,000 times, and data are conjured up. Furness (1978) stopped with biomass calculations and CIs, which seemed reasonable, while Roby et al (2003) added diet proportions, energy density and average fish weight to calculate the number of prey items consumed. But without the raw data and calculations, there is really no way to validate or verify that the results have any basis in reality. The model was adopted because it provided a way to conjure up numbers consumed rather than it had any validity given the way it is used. The simulation approach adopted might work with fewer categories to avoid small proportion problems, sufficient data to populate the tables with valid means and SEs, and fewer calculations to minimize propagating error, but this is unlikely given the admittedly small sample sizes.

The respondent should read the papers referenced. Again the methods and protocol ARE NOT detailed in the referenced papers although Furness (1978) has more detail than Roby et al (2003). Same is true for subsequent references back to the original papers in yearly reports. If details would have been included and reviewers knew a little stats, Roby et al 2003 likely would have been rejected simply because of the manner in which “random” numbers are manipulated if the latter had been revealed. DCCO yearly reports don’t mention a distribution although “random normal” was in Furness (1978) and Roby et al (2003). Based on responses in FEIS and examination of data tables in Tillamook Bay 2012 simulation, there was an attempt to use random normal distributions with considerable fudging of the distributions as an afterthought or from the outset? In any case, the approach is a flawed. Not only are the “random normal” numbers fudged by forcing them to sum to 1 for each iteration, now there is the admission that the distributions were truncated supposedly by setting negative values to zero. Apparently negative values were simply discarded and iterations were run until there were 1,000 positive values or some other internal adjustment was made. Depending on how many times this was done, this would shift “best estimates” to the right, or increase them. Since the diet proportions are relative, this would decrease other proportions. In science there should be a Methods section that allows other researchers to repeat the experiment, or simulation in this case, to verify results. Not only was this not done, now the code is being deemed proprietary and the raw data appear unavailable or secret. This requires both the computer code for inspection to see how many fudge factors were incorporated as well as the raw input data to determine if data were sufficiently robust. Dribbling out details in piece meal and obstructionist fashion is not science.

Regarding the respondent’s second paragraph on adjusting the distribution, when you start fiddling with the “simulated” data, one wonders what else was adjusted in the simulations and the respondent’s concept of reality, biological or otherwise. The respondent’s assertion that adjustments were done for biological reality is silly – it is a simulation not reality. Even modelers now recognized the futility is attempting to simulate reality and the problem of propagating error. For example, from Ford et al. (2010), *Energetic modelling: A comparison of the different approaches used in seabirds. Comparative Biochemistry and Physiology, Part A* 158 (2011) 358–365 “A majority of mathematical models applied to ecology tend to take into account as many parameters affecting output variables as possible, thereby seeking the closest-possible match with a natural situation. However, such an increase in the number of considered variables

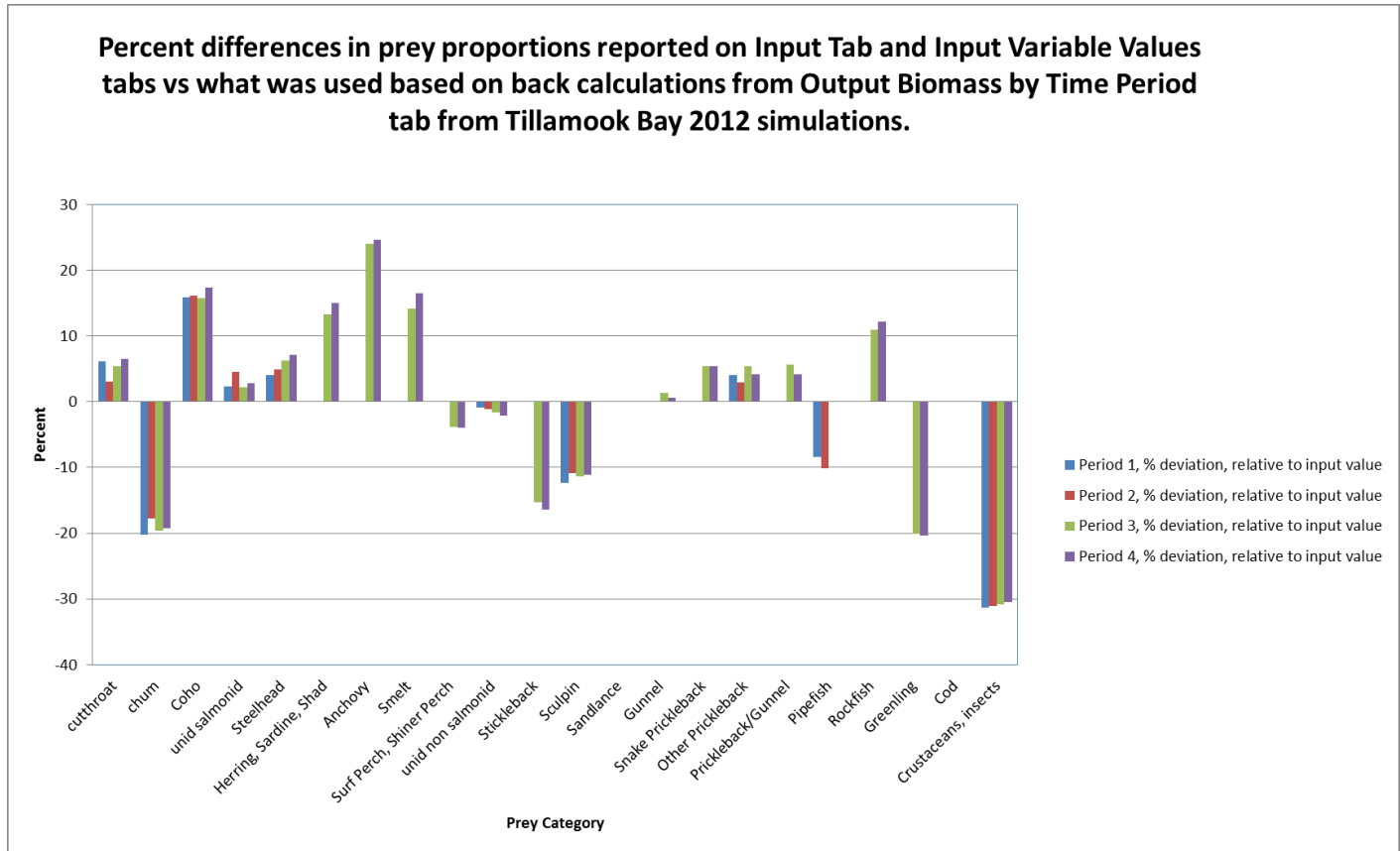
simultaneously accumulates the uncertainty associated to each one of them, and might thereby strongly increase the overall, cumulative variance of the model output.” In other words, as more variables are added, especially those with relatively large errors, the errors propagate through the calculations resulting in a cumulative error that is greater than a simple summation. This is simply a mathematical certainty; there is no “might” about it.

Given minuscule sample sizes for diet proportions there is really no attempt to represent biological reality. Just stick some data in and generate results including as many variables as possible while manipulating error to make the results seem reasonable. Regarding biological reality and getting every iteration to sum to one, how was the adjustment done? By adding/subtraction small values, recalculating each proportion as a proportion of the total summed proportion for the iteration? The respondent doesn't appear to grasp the concept of random and normal and the propagating error in this kind of fiddling goes through the roof. Finally, the respondent fails to realize that the starting seed values used to generate the distributions after sufficient iterations, whether it is 1,000, 10,000 or a million, values (“means”) computed from the iterations will be the same as the starting seed values assuming your numbers generator is working. In other words they will sum to 1. The point of random distribution is to generate a random normal distribution for each value, especially if that is what is implied or stated. The diet proportions are related because they are proportions of a diet but this is a problem inherent in the Roby et al (2003) modification of Furness (1978). A related feature of constraining the distribution as was done is that CIs are underestimated.

S-5. Respondent is simply making this up or doesn't understand what was done. If the distribution was “truncated normal” or binomial, the CIs would be asymmetrical and they are not. From plots of the values, it looks like what was done was back cast or simulate an SE that would result in few negative values then generate a symmetrical distribution using that value. Due to the truncation on the left tail, this would underestimate the dispersion of the right tail resulting gross underestimation of CIs. What was really done?

“Qualitative similar” is an oxymoron in science. Don Lyons in pers. com. also noted that the bioenergetics runs were qualitatively similar to PIT tag data, which essentially means DCCOs eat salmon. A lot of room there for quantitatively different.

S-6. Again the respondent is clueless. The values and calculation come directly from the Tillamook Bay 2012 worksheet generated from what we now know is the same OSU/BRNW model. In that run, Periods 1 & 2 and Periods 3-4 diet proportions were pooled probably because a lack of data. E.g., only two stomach in Period 1. Then these were applied as if independent samples to Periods 1, 2, 3 & 4 by those running the numbers. So these are the proportions input to the model. Biomass is projected as detailed in my comments resulting in total biomass for periods which is apportioned using the pooled proportions by design. So basically what the respondent is suggesting is that $4 \times 2 = 8$, but $8/2 = 4.3$. So this problem still stands. The respondent may also not realize the difference between a proportion and a percent? (To convert proportion to percent, move the decimal two places to the right.) Because of the exorbitant extrapolations, a difference of 1% (or proportional difference of .01) can change numbers significantly especially because of the interrelated nature of the proportions. I redid the figure (below) by computing % differences for each period, and the error in calculations is more striking. Despite assurance that it doesn't matter, there are errors in the code, fudge factors, corrections to numbers using hidden rules or protocols, or all of these. Only by examining the code can these bugs or errors be tracked down.



S-7. Respondent is rehashing what I already discovered as the “hidden protocol” and detailed in the original comments. I’d challenge anyone to stand up in front of a crowd of scientists and explain the ludicrous nature of transforming binary data to proportions significant to three decimal places. Note the mistake in calculation of chum in Tillamook Bay 2012 calculations resulted in doubling the estimate. The genetic id method is impressive, but there needs to be additional work on the quantitative aspect before it is employed in any management.

S-8. Now the Tillamook Bay numbers are deemed preliminary although the results were titled “Final Report” and results were submitted to the Oregon legislature to support cormorant control. Input data for ESI please. How many errors were made in those calculations? There were so few fish in the Tillamook analysis (~29) that the analysis is pointless which would have been evident if details from ESI simulation had been published or reviewed prior to using the method for 10+ years.

S-9. There is no more here than I summarized in my Table 5 & Fig 5 from my original comments except for some additional caveats and excuses and references back to other presentation of summary data that provide little additional detail. For example, from Chap 1, page 11: “Efforts to validate these assumptions were made whenever possible, and these and other assumptions, caveats, discussion points, and sample sizes are presented in more detail in Appendix C and in Roby et al. (2003) and Lyons (2010).” Not so. One notable change or omission is that the minuscule sample sizes now result from a concern for animal welfare, population impacts, or permit issues. Missing are previous references to Trites & Joy (2005, Dietary analysis from fecal samples: how many scats are enough? J.

Mammalogy 86:704-712) that was used to justify the small samples of approximately 15 stomachs per sample period at East Sand Island. In previous comments I noted the too small sample sizes and that Trites & Joy did not apply to what was being attempted for DCCOs stomach analysis.

From Table 5, gm of salmon in the sample ranged from ~2,000 to 4,700 over five years. Apportioned into 20-30 prey types with only 15 stomachs for 10+ sample periods, with 0-4 prey types per stomach, and the data tables are extremely sparse (i.e., mostly zeros). Then these are used to compute means and SDs, later converted to SEs or simulated, after first computing percent (or proportions) of each prey type per stomach. I provided an example from Tillamook Bay to illustrate the data table. In that run, problem with stomach sampling included over parsing of data and small samples were obvious from examination of the data matrix is that diet proportional data were based on a preponderance of zeroes. In the non-salmonid and overall salmonid data matrix (page 10), in some cases means and SDs for prey categories were based on zeroes except for a single stomach with prey (6 of 21 instances) or zeroes with fewer than four stomachs with prey (12 of 21 instances instances). Although these sparse and non-normal data clearly violate assumptions of normality, were assumed to be normal and used as estimators of parametric means, SDs, and SEs for input for the model.

I repeated Table 5 and Fig. 5 from the original comments. Data are directly from Bird Research NW summary reports. Note the column titled “number of fish, *minimum in sample”. These are salmonids or salmonid detections in five respective years, were 188, 249, 332, 260, and 137 salmonids. Assuming these were parsed across 10 periods, the number in each sample period ranges from 14-33 salmonids. Parsed in to five salmonid prey types and the data tables consist mostly of zero cells. This then is similar to the Tillamook Bay calculations where an estimated 29 salmonids were extrapolated to 51,000 by mistake, then corrected to 29,000. However the extrapolations used at East Sand Island were more extreme. The extent of the extrapolation are in the column labeled “1 fish=” and in Fig. 5 where one steelhead, which seems to be the poster species for the control effort, extrapolates to 1/4 to 1/3 of a million consumed. I would agree that data are minimal noting the additional comment in Appendix C.

Table 5. Data copy/pasted and parsed (gray cells) from Bird Research NW website (<http://www.birdresearchnw.org/project-info/weekly-update/columbia-river-estuary>) yearly breakdown of consumption. Using the total gm of fish in yearly samples, the gm of salmonids can be estimated, followed by gm of species or age, divided by average size of fish yields the estimated number of fish in original sample. Dividing millions of fish from Roby et al 2009-2013 by estimated number in original sample yields the extrapolation "rate" for each fish (Fig. 5).

Year	gm identifiable fish	# stomachs	% salmonids	gm salmonids	92% id'd	gm of salmonid	proportion	# fish, minimum* in sample "=ColK/ColO"	Extrapolated millions of salmonids from Roby et al 2009-2013 see ColB	average gm/fish, from Tillamook Bay 2012 workbook, or Lyons et al	gm/stomach	millions/ # of fish	1 fish =
2013	25,281	134	11%	2,781	92% id'd								
	juvenile salmonids comprised 10.7% of the diet												
	11.4 million sub-yearling Chinook smolts (95% c.i. = 6.9 – 15.9 million; 70% of total)												
						1,947	0.700	162	11.4	12		0.070	70,275
	2.7 million coho salmon smolts (95% c.i. = 2.0 – 3.4 million; 17% of total smolt consumption),												
						473	0.170	16	2.7	29.1		0.166	166,196
	1.0 million steelhead smolts (95% c.i. = 0.8 – 1.3 million; 6%),												
						167	0.060	3	1	59.6		0.357	357,197
	0.9 million yearling Chinook salmon smolts (95% c.i. = 0.6 – 1.1 million; 5%),												
						139	0.050	5	0.9	25.7		0.166	166,348
	0.2 million sockeye salmon smolts (95% c.i. = 0.05 – 0.3 million; 1%; Figure 70).												
						28	0.010	1	0.2	19.3		0.139	138,803
						TOTAL		188	16		20.5	0.086	86,117
2012	21,331	134	19%	4,053									
	18.9 million smolts (95% c.i. = 14.0 – 23.8 million),												
	10.8 million smolts or 57% were sub-yearling Chinook salmon (95% c.i. = 6.8 – 14.8 million),												
						2,310	0.570	193	10.8	12		0.056	56,100
	4.8 million smolts or 26% were coho salmon (95% c.i. = 3.5 – 6.0 million),												
						1,054	0.260	36	4.8	29.1		0.133	132,555
	1.7 million smolts or 9% were steelhead (95% c.i. = 1.3 – 2.1 million),												
						365	0.090	6	1.7	59.6		0.278	277,772
	1.5 million smolts or 8% were yearling Chinook salmon (95% c.i. = 1.0 – 2.0 million), and												
	0.1 million smolts or 0.6% were sockeye salmon (95% c.i. = 0.00 – 0.3 million);												
						324	0.080	13	1.5	25.7		0.119	118,897
						24	0.006	1	0.1	19.3		0.079	79,367
						TOTAL		249	19		30.4	0.076	75,989
2011	24,788	135	19%	4,710									
	76% were sub-yearling Chinook salmon (best estimate = 15.6 million smolts; 95% c.i. = 10.6 – 20.2 million),												
						3,579	0.760	298	15.6	12		0.052	52,299
	13% were coho salmon (best estimate = 2.7 million smolts; 95% c.i. = 2.1 – 3.4 million),												
						612	0.130	21	2.7	29.1		0.128	128,327
	6% were steelhead (best estimate = 1.2 million smolts; 95% c.i. = 0.9 – 1.4 million),												
						283	0.060	5	1.2	59.6		0.253	253,094
	4% were yearling Chinook salmon (best estimate = 0.9 million smolts; 95% c.i. = 0.7 – 1.1 million),												
						188	0.040	7	0.9	25.7		0.123	122,778
	0.4% were sockeye salmon (best estimate = 0.01 million smolts; 95% c.i. = 0.00 – 0.05 million)												
						19	0.004	1	0.01	19.3		0.010	10,245
						TOTAL		332	20		34.7	0.061	61,407
2010	23,356	134	16.50%	3,854									
	2010 was 19.2 million smolts (95% c.i. = 14.6 – 23.8 million),												
	69.8% were sub-yearling Chinook salmon (best estimate = 13.4 million; 95% c.i. = 9.1 – 17.6 million),												
						2,690	0.698	224	13.4	12		0.060	59,779
	15.6% were coho salmon (best estimate = 3.0 million; 95% c.i. = 2.3 – 3.7 million),												
						601	0.156	21	3	29.1		0.145	145,214
	7.8% were steelhead (best estimate = 1.5 million; 95% c.i. = 1.2 – 1.8 million),												
						301	0.078	5	1.5	59.6		0.297	297,413
	6.8% were yearling Chinook salmon (best estimate = 1.3 million; 95% c.i. = 1.0 – 1.6 million),												
						262	0.068	10	1.3	25.7		0.127	127,493
	and 0.2% were sockeye salmon (best estimate = 0.03 million; 95% c.i. = 0.01 – 0.06 million);												
						8	0.002	0.4	0.03	19.3		0.075	75,122
						TOTAL		260	19		28.8	0.074	73,831
2009	21,830	133	9%	1,965									
	74% were sub-yearling Chinook salmon (best estimate = 8.3 million; 95% c.i. = 5.1 – 11.4 million)												
						1,454	0.740	121	8.3	12		0.069	68,506
	12% were coho salmon (best estimate = 1.4 million; 95% c.i. = 1.0 – 1.7 million),												
						236	0.120	8	1.4	29.1		0.173	172,800
	7% were steelhead (best estimate = 0.8 million; 95% c.i. = 0.6 – 1.0 million),												
						138	0.070	2	0.8	59.6		0.347	346,691
	6% were yearling Chinook salmon (best estimate = 0.7 million; 95% c.i. = 0.5 – 0.8 million), and												
						118	0.060	5	0.7	25.7		0.153	152,610
	< 1% were sockeye salmon (best estimate = 0.02 million; 95% c.i. = 0.01 – 0.03 million; Figure 32)												
						18	0.009	1	0.02	19.3		0.022	21,830
						TOTAL		137	11		14.8	0.082	81,857

*from foregut, so fish mostly whole

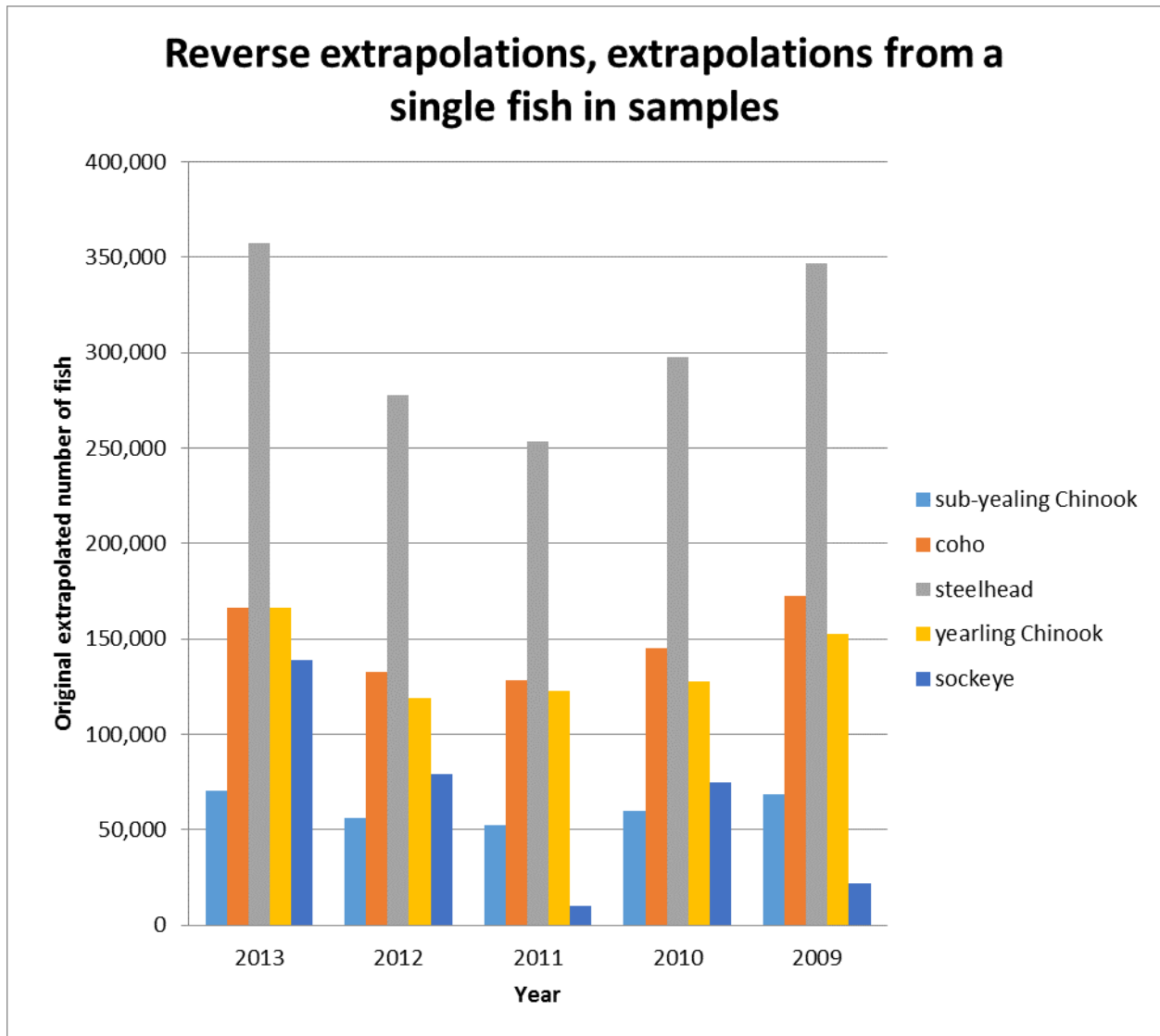


Figure 5. Computed from summary data from Roby et al. 2009-2013, see Table 5. Each fish in DCCO foregut samples was extrapolated to the number on the y-axis illustrating the tenuous nature for original projection. Reverse extrapolations were necessitated in the absence of raw data.

Comments on supplemental comments from Fredricks, Appendix D on the amorphous 3.6% gap

Fredricks, 3.) Simplicity/availability. While the OSU bioenergetics model and the NOAA smolt population estimates are far from simple, they were readily available in the literature and were easily adapted to the need and timeframe of the BiOp analysis. The PIT tag data that would fit the smolt population as a whole were not readily available in the literature and an ESU by ESU analysis was

tainted by the complications of tagging rates per ESU and the ultimate issue of how one would manage birds for any particular fish stock. Final Environmental

Response from G. W. Shugart. This is untrue on two points. A better title would be "complex/unavailable" or was the title facetious? First the model is not available in the literature and a view into what is essentially a black box was possible through use in Tillamook Bay 2012, which I provided a brief overview in the original comments. Secondly, Fredricks is not dealing with actual data with OSU/BRNW simulations based on bioenergetics, but exorbitant extrapolations from few data, presumably similar to the Tillamook Bay 2012 simulation. See Table 5 and Fig 5 from the original comment repeated above. Like the simulation, the raw input data to the simulations are unavailable. Without the code or input there is no way to evaluate the simulated consumption numbers being used for extrapolations that Fredricks is doing with extrapolations beyond the initial extrapolations. The extrapolation could be reasonable or simply garbage depending on the robustness of the input data. To simplify Fredricks' analysis, there is no dispute that zero DCCOs eat zero salmon and a lot of DCCOs eat a lot of salmonids. So basically split the difference and to reduce the population to point where DCCOs consume less than a moderate amount. Fredricks "analysis" is similar to Good et al.'s in the Caspian Tern EIS with extrapolations upon extrapolations on simulated data of simulated benefits in reduction in predation. Basically, make something up to rationalize management.

From Fredricks, Appendix D.

4.) Error. Both approaches have error associated with them. The bioenergetics consumption and cormorant population estimates had errors reported in the literature, while the smolt population estimates did not. The PIT tag consumption estimates had good precision owing to the known starting populations, however, there were significant uncertainties associated with tag deposition and detection rates as well as the issue erratic tagging per species and ESU.

G. W. Shugart response. Again the bioenergetics input data and associated error are not available and error was not properly assessed. Rather the CIs associated with the simulated data are simulated and presented to give the estimates legitimacy, but are underestimated based on the Tillamook Bay 2012 simulations and the manipulations of distributions that were revealed in the FEIS. If done correctly the errors would likely be greater than the estimated means, making the data ridiculous. But for Fredricks' extrapolations, the error is ignored anyway so not clear what the point was. Note my point from page 1 that the only point of running the Monte Carlo simulation was to estimate error.

From Fredricks, Appendix D.

In summary, there is nothing inappropriate or incorrect about either approach, however it was my opinion that the bioenergetics/smolt population approach was a better fit for the supplemental Biological Opinion. This approach was discussed with the bird researchers prior to the development of the final BiOp and they did not believe that it was inappropriate for the intended use. Regardless of which approach was used in the 2014 Supplemental BiOp, it is appropriate to use all lines of evidence when developing and assessing the cormorant management plan.

G. W. Shugart response. It is clearly inappropriate to use miniscule samples, simulated data replete with numerous manipulation, and simulated consumption numbers to create an allusion of science.

The theme of the management effort is that predators are inherently harmful to the resource so it really doesn't matter what was quantified.