Validation of sea lice data extracted from Wild Juvenile Salmonid Monitoring Program PDF reports

C. Mimeault

June 13, 2021

Some companies operating finfish aquaculture sites on the West Coast of Canada contract third parties to conduct juvenile wild salmon sea lice monitoring in British Columbia (BC). All reports can be found here in PDF format.

As of June 2021, reports were available for surveys conducted from 2014 to 2020 in five different regions of BC. Sea lice counts at the fish level are included in the Appendix III of the following reports:

- Broughton Archipelago (2016-2020): 5 years
- Discovery Islands (2017-2020): 4 years
- Quatsino (2016-2020): 5 years

Data extraction from PDFs

When available, fish-level data were extracted from the PDF documents using Camelot and csvkit librairies using bash code available here. The Rmarkdown file to generate the present document is available here.

Validation of extracted data

Extracted data were validated by reproducing some tables included in the reports. Note that not all sampled fish were collected. Only tables using data related to Appendix III could be reproduced. The relevant tables from the reports are reproduced in this document and compared to the results using extracted data. In case of discrepancies, reasons were explored.

Broughton Archipelago reports

2016

The 2016 Wild Juvenile Salmonid Monitoring Program Broughton Archipelago, BC report can be found here. The report includes, among other things, sea lice counts on chum salmon (CM), coho salmon (CO), Chinook salmon (CH), pink salmon (PK) and threespine stickleback (TSB) collected at 32 beach seine sites in 2016 in Broughton Archipelago.

Table 3

Table 3 as included in the 2016 Broughton Archipelago report:

Table 3: The total of collected individuals of each fish species captured in the Broughton Archipelago, BC in April and May 2016, and the percentage of the total capture population that they represent.

Common Name	Capture Totals (% of total capture population)	Collection Totals	Collection %
pink salmon	1451 (58.9%)	430	29.6
chum salmon	984 (40.0%)	512	52.0
coho salmon	25 (1.0%)	25	100
Three spine stickleback	2 (0.1%)	2	100
All species	**2462**	**969**	**39.4**

Figure 1: Table 3 from the 2016 Broughton Archipelago PDF document

Table 3 reproduced using extracted data:

fish_species	n
pink salmon	430
chum salmon	512
coho salmon	25
chinook salmon	0
threespine stickleback	2

The total number of individuals collected per species in Table 3 of the PDF and using the extracted data are the same. Additionally, based on the extracted data, there were 969 fish collected which also corresponds to collection totals in Table 3. Successfully reproducing Table 3 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 4

Table 4 as included in the 2016 Broughton Archipelago report:

Site	Site Name	Chum		Coho		Pink		Capture Total	Sample Total
		Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total		
1	Swanson Island Fish Farm	0	-	0	-	0	-	0	-
2	Midsummer Island Fish Farm (Potts Bay)	0	-	0	-	0	-	0	-
3	Chop Bay	0	-	0	-	4	4	4	4
4	Lady Island	51	36	0	-	39	37	90	73
5	Doctor Island Fish Farm	5	5	0	-	9	9	14	14
6	Brent Bay	0	-	0	-	0	-	0	-
7	Shelterless Bay	37	29	1	1	24	24	62	54
8	Lance Bay	0	-	2	2	0	-	2	2
9	Sargeaunt Pass	71	51	14	14	74	49	159	114
10	Humphrey Rock	70	30	0	-	670	30	740	60
11	Pumish Point	0	-	1	1	0	-	1	1
12	Oline Point	0	-	0	-	0	-	0	-
13	London Point	0	-	0	-	0	-	0	-
14	Miller Point	9	9	2	2	52	30	63	41
15	Kwatsi Point	17	17	0	-	61	31	78	48
16	Glacier Falls Fish Farm	3	3	0	-	8	8	11	11
17	Viner Sound	6	6	0	-	1	1	7	7
18	Denham Island	1	1	0	-	3	3	4	4
19	Baker Island	45	34	0	-	70	29	115	63
20	Jumper Island	29	29	0	-	37	26	66	55
21	Arthur Point	40	37	0	-	105	23	145	60
22	Wicklow Bay	4	4	0	-	3	3	7	7
A	Bennett Point Fish Farm	0	-	0	-	0	-	0	-
B	Sambo Point	0	-	0	-	0	-	0	-
C	Penphrase Passage	131	66	0	-	83	23	214	89
D	Harry Bay	0	-	0	-	0	-	0	-
E	Phillip Point West	42	36	3	3	0	-	45	39
F	Sutlej North	278	61	1	1	34	34	313	96
G	Codrington Point	55	30	1	1	0	-	56	31
H	Wehlis Bay Fish Farm	90	28	0	-	174	66	264	94
I	Alder Bay	0	-	0	-	0	-	0	-
J	Poppelwell Point	0	-	0	-	0	-	0	-
	Total	**984**	**512**	**25**	**25**	**1451**	**430**	**2460**	**967**

Table 4 reproduced using extracted data:

site	n
1	0
2	0
3	4
4	73
5	14
6	0
7	54
8	2
9	114
10	60
11	1
12	0
13	0
14	41
15	48
16	11
17	7
18	4
19	63
20	55
21	60
22	7
A	0
B	0
C	89
D	0
E	39
F	96
G	31
H	94
I	0
J	0

The total number of individuals collected per site in Table 4 of the PDF and using the extracted data are the same. Successfully reproducing Table 4 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 9

Table 9 as included in the 2016 Broughton Archipelago report:

Table 9: The number of sea lice in each life stage by species identified on the chum salmon sample population from the Broughton Archipelago in 2016. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 17/18/19, 2016	May 17/18/19, 2016
LEP Co	11	5
LEP C1	6	3
LEP C2	10	2
LEP C3	12	2
LEP C4	17	8
LEP PAM	5	3
LEP PAF	3	1
LEP AM	1	5
LEP AF	0	4
TOTAL LEP	**65**	**33**
CAL Co	5	2
CAL C1	52	59
CAL C2	8	7
CAL C3	4	4
CAL C4	8	3
CAL PAM	0	0
CAL PAF	0	0
CAL AM	0	3
CAL AF	0	9
TOTAL CAL	**77**	**87**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 2: Table 9 from the 2016 Broughton Archipelago PDF document

Table 9 reproduced using extracted data:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	11	5
lep_c1	6	3
lep_c2	10	2
lep_c3	12	2
lep_c4	17	8
lep_pam	5	3
lep_paf	3	1
lep_am	1	5
lep_af	0	4
lep_total	65	33
cal_co	5	2
cal_c1	52	59
cal_c2	8	7
cal_c3	4	4
cal_c4	8	3
cal_pam	0	0
cal_paf	0	0
cal_am	0	3
cal_af	0	9
cal_total	77	87

The sea lice counts on coho salmon collected in 2016 in Table 9 of the PDF and using the extracted data are the same (one NA value had to be dropped but it did not affect the results). Successfully reproducing Table 9 validates the extracted data used to generate this table.

Table 11

Table 11 as included in the 2016 Broughton Archipelago report:

Table 11: The number of sea lice in each life stage by species identified on the pink salmon sample population from the Broughton Archipelago in 2016. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 17/18/19, 2016	May 17/18/19, 2016
LEP Co	6	5
LEP C1	4	1
LEP C2	12	0
LEP C3	25	2
LEP C4	24	0
LEP PAM	6	1
LEP PAF	0	2
LEP AM	0	7
LEP AF	0	8
TOTAL LEP	**77**	**25**
CAL Co	1	0
CAL C1	40	34
CAL C2	24	2
CAL C3	15	1
CAL C4	6	0
CAL PAM	0	0
CAL PAF	0	0
CAL AM	1	4
CAL AF	0	12
TOTAL CAL	**87**	**53**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 3: Table 11 from the 2016 Broughton Archipelago PDF document

Table 11 reproduced using extracted data:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	6	5
lep_c1	4	1
lep_c2	12	0
lep_c3	25	1
lep_c4	24	0
lep_pam	6	1
lep_paf	0	2
lep_am	0	7
lep_af	0	8
lep_total	77	25
cal_co	1	0
cal_c1	40	34
cal_c2	24	2
cal_c3	15	1
cal_c4	6	0
cal_pam	0	0
cal_paf	0	0
cal_am	1	4
cal_af	0	12
cal_total	87	53

The sea lice counts on pink salmon collected in 2016 in Table 11 of the PDF and using the extracted data are NOT the same. There is a difference for the number of sea lice in LEP C3 in May (2 in the report vs 1 in the table reproduced with the extracted data). The error is probably not coming from the conversion of PDF into data as TOTAL LEP in May should add up to 26 with the numbers included in the table rather than to 25 as in the table and as in reproduced table. To verify, I referred to the data included in Appendix III of the PDF (using the PDF rather that the extracted data). There is only one instance of LEP C3 on pink salmon in May 2016 (on the 17th at Lady Island, see page 59). In other words, the results in Table 11 reproduced using extracted data are representative of the data in the Appendix III of the PDF. The discrepancy is explained. Successfully reproducing Table 11 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 13

Table 13 as included in the 2016 Broughton Archipelago report:

Table 13: The number of sea lice in each life stage by species identified on coho salmon from the Broughton Archipelago in 2016. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 17/18/19, 2016	May 17/18/19, 2016
LEP Co	0	0
LEP C1	0	0
LEP C2	0	0
LEP C3	0	0
LEP C4	0	0
LEP PAM	0	1
LEP PAF	0	2
LEP AM	0	2
LEP AF	0	5
TOTAL LEP	**0**	**10**
CAL Co	0	0
CAL C1	0	4
CAL C2	0	3
CAL C3	0	1
CAL C4	0	0
CAL PAM	0	0
CAL PAF	0	2
CAL AM	0	0
CAL AF	0	4
TOTAL CAL	**0**	**14**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 4: Table 13 from the 2016 Broughton Archipelago PDF document

Table 13 reproduced using extracted data:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	0	0
lep_c1	0	0
lep_c2	0	0
lep_c3	0	0
lep_c4	0	0
lep_pam	0	1
lep_paf	0	2
lep_am	0	2
lep_af	0	5
lep_total	0	10
cal_co	0	0
cal_c1	0	4
cal_c2	0	3
cal_c3	0	1
cal_c4	0	0
cal_pam	0	0
cal_paf	0	2
cal_am	0	0
cal_af	0	4
cal_total	0	14

The sea lice counts on coho salmon collected in 2016 in Table 13 of the PDF and using the extracted data are the same. Successfully reproducing Table 13 validates the extracted data used to generate this table.

Conclusions

The data extracted from PDF were used to reproduce the following tables in the report:

- Table 3: successfully reproduced indicating that related data validated
- Table 4: successfully reproduced indicating that related data validated
- Table 9: successfully reproduced indicating that related data validated
- Table 11: successfully reproduced indicating that related data validated
- Table 13: successfully reproduced indicating that related data validat

It is reasonable to conclude that the data extracted from the 2016 Broughton Archipelago report is representative of the data in the Appendix III of the report.

Discovery Islands reports

2017

The 2017 Wild Juvenile Salmonid Monitoring Program Broughton Archipelago, BC report can be found here. The report includes, among other things, sea lice counts on chum salmon (CM), coho salmon (CO), Chinook salmon (CH), pink salmon (PK) and threespine stickleback (TSB) sampled from 29 beach seine sites in 2017 in Broughton Archipelago.

Table 4

Table 4 as included in the 2017 Discovery Islands report:

Table 4: The total of collected individuals of each fish species captured in the Discovery Islands, BC in April and May 2017, and the percentage of the total capture population that they represent.

Common Name	Capture Totals (% of total capture population)	Collection Totals	Collection %
chum salmon	3703 (70.6 %)	942	25.4
pink salmon	1415 (27.0 %)	374	26.4
coho salmon	99 (1.9 %)	88	88.9
chinook salmon	26 (0.5 %)	26	100
threespine stickleback	1 (0.02 %)	1	100
All species	**5244**	**1431**	**27.3**

Figure 5: Table 4 from the 2017 Discovery Islands PDF document

"Collection Totals" column from Table 4 reproduced using extracted data:

fish_species	n
chum salmon	942
pink salmon	374
coho salmon	88
chinook salmon	26
threespine stickleback	1

The total number of individuals collected per species in Table 4 of the PDF and using the extracted data are the same. Additionally, based on the extracted data, there were 1431 fish collected which also corresponds to collection totals in Table 4. Successfully reproducing Table 4 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 7

Table 7 as included in the 2017 Discovery Islands report:

Site Location	Site Name	Pink		Chum		Coho		Chinook		Threespine stickleback		Capture Total	Sample Total
		Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total		
Pre-Exposure	Francisco Point	54	32	65	41	10	10	4	4	0	0	133	87
	Marina Island	29	29	98	32	7	7	6	6	0	0	140	74
	Rebecca Spit	6	6	74	29	17	17	0	0	0	0	97	52
	Viner Point	0	0	0	0	0	0	0	0	0	0	0	0
	SE Hill Island	2	2	21	21	0	0	0	0	0	0	23	23
	Penn Island	0	0	46	30	0	0	0	0	0	0	46	30
	Deepwater Bay	82	28	91	62	10	10	2	2	0	0	185	102
Pre-Exposure Site Totals		**173**	**97**	**395**	**215**	**44**	**44**	**12**	**12**	**0**	**0**	**624**	**368**
Post-Exposure	Raza	1	1	212	60	3	3	0	0	0	0	216	64
	Raza North	0	0	73	33	0	0	0	0	0	0	73	33
	Okisollo	1	1	235	60	0	0	0	0	0	0	236	61
	Owen Bay	0	0	23	23	0	0	0	0	0	0	23	23
	Rock Bay	5	5	270	30	0	0	1	1	0	0	276	36
	Discovery	92	32	131	60	0	0	0	0	0	0	223	92
	Nodales	605	60	547	47	0	0	0	0	0	0	1152	107
	Shoal Bay	8	8	285	35	0	0	0	0	0	0	293	43
	Fanny Bay	2	2	74	34	0	0	11	11	0	0	87	47
	Bickley Bay	3	3	86	30	0	0	0	0	0	0	89	33
	Cordero	0	0	3	3	0	0	0	0	0	0	3	3
	Knox Bay	15	15	82	47	0	0	1	1	0	0	97	63
	Bear Bay	321	41	110	50	0	0	0	0	0	0	431	91
	Chancellor Channel	0	0	6	6	0	0	0	0	0	0	6	6
	Race Passage	12	12	37	30	0	0	1	1	0	0	50	43
	Wellbore Channel	1	1	4	4	0	0	0	0	0	0	5	5
	Bessborough Bay	121	41	1000	60	0	0	0	0	0	0	1121	101
	Sunderland	10	10	41	32	1	1	0	0	0	0	52	43
	Blenkinsop Bay	7	7	20	20	1	1	0	0	1	1	29	29
	Primary 3	23	23	33	33	38	30	0	0	0	0	94	86
	Primary 1	15	15	36	30	12	9	0	0	0	0	63	54
	Beautiful Bay	0	0	0	0	0	0	0	0	0	0	0	0
Post Exposure Site Totals		**1242**	**277**	**3308**	**727**	**55**	**44**	**14**	**14**	**1**	**1**	**4620**	**1063**
Total Capture Totals		**1415**	**374**	**3703**	**942**	**99**	**88**	**26**	**26**	**1**	**1**	**5244**	**1431**

Figure 6: Table 7 from the 2017 Discovery Islands PDF document representing the number of individual fish (chum, coho and pink only) collected in the Broughton Archipelago in April and May 2016

"Sample Total" column from Table 7 reproduced using extracted data:

location	n
Francisco Point	87
Marina Island	74
Rebecca Spit	52
Viner Point	0
SE Hill Island	23
Penn Island	30
Deepwater Bay	102
Raza	64
Raza North	33
Okisollo	61
Owen Bay	23
Rock Bay	36
Discovery	92
Nodales	107
Shoal Bay	43
Fanny Bay	47
Bickley Bay	33
Cordero	3
Knox Bay	63
Bear Bay	91
Chancellor Channel	6
Race Passage	43
Wellbore Channel	5
Bessborough Bay	101
Sunderland	43
Blenkinsop Bay	29
Primary 3	86
Primary 1	54
Beautiful Bay	0

The total number of individuals collected per location in Table 7 of the PDF and using the extracted data are the same. Successfully reproducing Table 7 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 17

Table 17 as included in the 2017 Discovery Islands report:

Table 17: The number of sea lice in each life stage by species identified on the Pre-Exposure chum salmon sample population from the Discovery Islands in 2017. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 9/10/11, 2017	May 23/25/26, 2017
LEP Co	0	2
LEP C1	0	0
LEP C2	0	3
LEP C3	0	2
LEP C4	0	0
LEP PAM	0	1
LEP PAF	0	0
LEP AM	0	0
LEP AF	0	0
TOTAL LEP	**0**	**8**
CAL Co	25	0
CAL C1	47	5
CAL C2	2	3
CAL C3	0	3
CAL C4	0	1
CAL PAM	0	0
CAL PAF	0	0
CAL AM	0	0
CAL AF	0	1
TOTAL CAL	**74**	**13**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 7: Table 17 from the 2017 Discovery Islands PDF document

Table 17 was reproduced using extracted data. To do so, first I filtered for chum salmon data and for the pre-exposure locations, then I re-formatted the data (using gather), then calculated numbers of sea lice for April, then for May. Finally, I brought the April and May data together.

Table 17 reproduced using extracted data:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	0	0
lep_c1	0	2
lep_c2	0	3
lep_c3	0	2
lep_c4	0	0
lep_pam	0	1
lep_paf	0	0
lep_am	0	0
lep_af	0	0
lep_total	0	8
cal_co	25	0
cal_c1	47	5
cal_c2	2	3
cal_c3	0	3
cal_c4	0	1
cal_pam	0	0
cal_paf	0	0
cal_am	0	0
cal_af	0	1
cal_total	74	13

The sea lice counts on chum salmon collected at pre-exposure locations in May in Table 17 of the PDF and using the extracted data are NOT the same. There is a small discrepancy for the LEP Co and LEP C1 in May. To understand the discrepancy, I referred to the data included in Appendix III of the PDF (using the PDF rather that the extracted data):

For counts of LEP Co in chum salmon (CM) at pre-exposure locations (listed below) in May:

- Francisco Point (p. 87-88): 0
- Marina Island (p.88): 0
- Rebecca Spit (p.91-92): 0
- Viner Point - no samples
- SE Hill Island (p.92-93): 0
- Penn Island (p. 89-90): 0
- Deepwater Bay (p. 86-87): 0

Based on the results included in the Appendix III of the PDF, there were 0 LEP Co observed on chum salmon captured at pre-exposure locations in May.

For counts of LEP C1 in chum salmon at pre-exposure locations in May:

- Francisco Point (p. 87-88): 1
- Marina Island (p.88): 0
- Rebecca Spit (p.91-92): 0
- Viner Point - no samples
- SE Hill Island (p.92-93): 0
- Penn Island (p. 89-90): 1
- Deepwater Bay (p. 86-87): 0

Based on the results included in the Appendix III of the PDF, there were 2 LEP C1 observed on chum salmon captured at pre-exposure locations in May.

In other words, the results in Table 17 reproduced using extracted data are representative of the data in the Appendix III of the PDF. The discrepancy is explained. Successfully reproducing Table 17 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 23

Table 23 as included in the 2017 Discovery Islands report:

Table 23: The number of sea lice in each life stage by species identified on the Post-Exposure chum salmon sample population from the Discovery Islands in 2017. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 9/10/11, 2017	May 23/25/26, 2017
LEP Co	4	2
LEP C1	2	0
LEP C2	0	1
LEP C3	1	5
LEP C4	0	4
LEP PAM	0	4
LEP PAF	0	0
LEP AM	0	0
LEP AF	0	0
TOTAL LEP	**7**	**16**
CAL Co	7	5
CAL C1	3	11
CAL C2	1	0
CAL C3	0	1
CAL C4	0	0
CAL PAM	0	0
CAL PAF	0	0
CAL AM	0	3
CAL AF	0	0
TOTAL CAL	**11**	**20**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 8: Table 23 from the 2017 Discovery Islands PDF document

Table 23 was reproduced using extracted data. To do so, first I filtered for chum salmon data and for the post-exposure locations, then I re-formatted the data (using gather), then calculated numbers of sea lice for April, then for May. Finally, I brought the April and May data together.

Table 23 reproduced using extracted data:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	4	2
lep_c1	2	0
lep_c2	0	1
lep_c3	1	5
lep_c4	0	4
lep_pam	0	4
lep_paf	0	0
lep_am	0	0
lep_af	0	0
lep_total	7	16
cal_co	7	5
cal_c1	3	11
cal_c2	1	0
cal_c3	0	1
cal_c4	0	0
cal_pam	0	0
cal_paf	0	0
cal_am	0	3
cal_af	0	0
cal_total	11	20

The sea lice counts on chum salmon collected at post-exposure locations in May in Table 23 of the PDF and using the extracted data are the same. Successfully reproducing Table 23 validates the extracted data used to generate this table.

Conclusions

The data extracted from PDF were used to reproduce the following tables in the report:

- Table 4 was successfully reproduced indicating that the related data are validated
- Table 7 was successfully reproduced indicating that the related data are validated
- Table 17 was successfully reproduced indicating that the related data are validated
- Table 23 was successfully reproduced indicating that the related data are validated

It is reasonable to conclude that the data extracted from the 2017 Discovery Islands report are representative of the data in the Appendix III of the report. The few discrepencies were explained and were not related to an error in the extracted data.

Quatsino reports

2016

The 2016 Wild Juvenile Salmonid Monitoring Program Quatsino Sound, BC report can be found here. The report includes, among other things, sea lice counts on chum salmon (CM), coho salmon (CO), Chinook salmon (CH) and pink salmon (PK) collected by beach seine in 2016 in Quatsino Sound and Holberg Inlet.

Table 3

Table 3 as included in the 2016 Quatsino report:

Table 3: The total of collected individuals of each fish species captured in Quatsino Sound and Holberg Inlet, BC in April and May 2016, and the percentage of the total capture population that they represent.

Common Name	Capture Totals (% of total capture population)	Collection Totals	Collection %
chum salmon	938 (97.7%)	235	25.1
coho salmon	1 (0.1%)	1	100
chinook salmon	19 (2.0%)	19	100
pink salmon	2 (0.2%)	2	100
All species	**960**	**257**	**26.8**

Figure 9: Table 3 from the 2016 Quatsino PDF document

Table 3 reproduced using extracted data:

fish_species	n
chum salmon	235
coho salmon	1
chinook salmon	19
pink salmon	2

The total number of individuals collected per species in Table 3 of the PDF and using the extracted data are the same. Additionally, based on the extracted data, there were 257 fish collected which also corresponds to collection totals in Table 3. Successfully reproducing Table 3 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 4

Table 4 as included in the 2016 Quatsino report:

Table 4: The number of captured fish (Capture Total) and the number of individual fish collected (Sample Total) from each of the 10 sample sites in Quatsino Sound and Holberg Inlet, BC in April and May 2016.

SITE	Chum		Coho		Pink		Chinook		Capture Total	Sample Total
	Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total	Capture Total	Sample Total		
1	64	44	0	0	0	0	8	8	72	52
2	243	33	0	0	1	1	6	6	250	40
3	61	30	0	0	0	0	4	4	65	34
4	0	0	0	0	0	0	0	0	0	0
5	8	8	0	0	0	0	0	0	8	8
6	0	0	0	0	0	0	0	0	0	0
7	33	30	1	1	0	0	0	0	34	31
8	0	0	0	0	0	0	0	0	0	0
9	457	60	0	0	1	1	1	1	459	62
10	72	30	0	0	0	0	0	0	72	30
Total	**938**	**235**	**1**	**1**	**2**	**2**	**19**	**19**	**960**	**257**

Figure 10: Table 4 from the 2016 Quatsino PDF document

Table 4 reproduced using extracted data:

site	n
Site 1	53
Site 2	40
Site 3	34
Site 4	0
Site 5	8
Site 6	0
Site 7	31
Site 8	0
Site 9	61
Site 10	30

The total number of individuals collected per site in Table 4 of the PDF and using the extracted data are NOT the same. They differ for the number of fish collected at Site 1 and Site 9.

To understand the discrepancy, I referred to the data included in Appendix III of the PDF (using the PDF rather that the extracted data):

For Site 1:

- April 5, 2016: 31 fish collected (see page 53);
- May 4, 2016: 17 fish collected (see page 39); and
- May 4, 2016: 5 fish collected (see page 40).

So a total of 53 fish collected at Site 1 as in Table 4 reproduced using extracted data.

For Site 9:

- April 5, 2016: 31 fish collected (see page 38);
- May 4, 2016: 26 fish collected (see page 40); and
- May 4, 2016: 4 fish collected (see page 41).

So a total of 61 fish collected at Site 9 as in Table 4 reproduced using extracted data.

The results in Table 4 reproduced using extracted data are representative of the data in the Appendix III of the PDF. The discrepancy is explained. Successfully reproducing Table 4 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 8

Table 8 as included in the 2016 Quatsino report:

Table 8: The number of sea lice in each life stage by species identified on the chum salmon sample population from Quatsino Sound and Holberg Inlet in 2016. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 5, 2016	May 4, 2016
LEP Co	2	1
LEP C1	1	3
LEP C2	1	1
LEP C3	0	7
LEP C4	0	2
LEP PAM	1	5
LEP PAF	0	0
LEP AM	1	0
LEP AF	0	0
TOTAL LEP	**6**	**19**
CAL Co	1	2
CAL C1	9	9
CAL C2	3	1
CAL C3	0	2
CAL C4	0	2
CAL PAM	0	0
CAL PAF	0	0
CAL AM	0	0
CAL AF	0	0
TOTAL CAL	**13**	**16**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 11: Table 8 from the 2016 Quatsino PDF document

Table 8 was reproduced using extracted data. To do so, first I filtered for chum salmon data, then I re-formated the data (using gather), then calculated numbers of sea lice for April, then for May. Finally, I brought the April and May data together.

Table 8 as included in the 2016 Quatsino report:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	2	1
lep_c1	1	3
lep_c2	1	1
lep_c3	0	7
lep_c4	0	2
lep_pam	1	5
lep_paf	0	0
lep_am	1	0
lep_af	0	0
lep_total	6	19
cal_co	1	2
cal_c1	9	9
cal_c2	3	1
cal_c3	0	2
cal_c4	0	2
cal_pam	0	0
cal_paf	0	0
cal_am	0	0
cal_af	0	0
cal_total	13	16

The number of sea lice per species and life stage in Table 8 of the PDF and using the extracted data are the same. Successfully reproducing Table 8 validates the extracted data used to generate this table. Moving on to the next relevant table.

Table 10

Table 10 as included in the 2016 Quatsino report:

Table 10: The number of sea lice in each life stage by species identified on chinook and pink salmon from Quatsino Sound and Holberg Inlet in 2016. LEP = *Lepeophtheirus salmonis* CAL = *Caligus clemensi*

Life Stage[1]	April 5, 2016	May 4, 2016
LEP Co	2	2
LEP C1	0	0
LEP C2	0	1
LEP C3	0	0
LEP C4	0	1
LEP PAM	0	0
LEP PAF	0	0
LEP AM	0	0
LEP AF	0	0
TOTAL LEP	**2**	**4**
CAL Co	0	0
CAL C1	0	9
CAL C2	0	0
CAL C3	0	0
CAL C4	0	0
CAL PAM	0	0
CAL PAF	0	0
CAL AM	0	0
CAL AF	0	0
TOTAL CAL	**0**	**9**

[1] Lice life stage codes: Co = copepodid, C1-4 = chalimus 1-4, PAM = pre-adult male, PAF = pre-adult female, AM = adult male, AF = adult female.

Figure 12: Table 10 from the 2016 Quatsino PDF document

Table 10 was reproduced using extracted data. To do so, first I filtered for chinook and pink salmon data, then I re-formatted the data (using gather), then calculated numbers of sea lice for April, then for May. Finally, I brought the April and May data together.

Table 10 as included in the 2016 Quatsino report:

sea_lice_species_and_life_stages	lice_counted_april	lice_counted_may
lep_co	2	2
lep_c1	0	0
lep_c2	0	1
lep_c3	0	0
lep_c4	0	1
lep_pam	0	0
lep_paf	0	0
lep_am	0	0
lep_af	0	0
lep_total	2	4
cal_co	0	0
cal_c1	0	9
cal_c2	0	0
cal_c3	0	0
cal_c4	0	0
cal_pam	0	0
cal_paf	0	0
cal_am	0	0
cal_af	0	0
cal_total	0	9

The number of sea lice per species and life stage in Table 10 of the PDF and using the extracted data are the same. Successfully reproducing Table 10 validates the extracted data used to generate this table.

Conclusions

The data extracted from PDF were used to reproduce the following tables in the report:

- Table 3: successfully reproduced indicating that related data validated
- Table 6: successfully reproduced indicating that related data validated
- Table 8: successfully reproduced indicating that related data validated
- Table 10: successfully reproduced indicating that related data validated

It is reasonable to conclude that the data extracted from the 2016 Quatsino report is representative of the data in the Appendix III of the report.

Conclusion for validation of the extracted data

None of the discrepancies were due to errors in the extracted data. I therefore concluded that is reasonable to assume that the extracted data represent the data in the Appendix and will therefore move on to the standardization of the report.

Standardization of datasets

Collation of datasets

Re-validation of dataset

Final dataset