Report on 2007 tailed frog monitoring results, Flathead and Yahk Rivers, Nelson Forest Region, near Cranbrook, BC.

MoF Contract CECO81

Report to

Kathy Paige
Biodiversity Branch
Environmental Stewardship Division
Ministry of Water, Land and Air Protection
PO Box 9398 Stn Prov Govt
Victoria, BC
V8W 9M9

By

Cordilleran Geoscience,
Box 612, Squamish BC
V8B 0A5

Table of Contents

Appendices

1.0 Introduction

A Rocky Mountain tailed frog (*Ascaphus montanus*) monitoring pilot study was conducted in 2005 (Friele and Dupuis 2006a). The purpose of the pilot was to test the feasibility of the monitoring protocols proposed by Dupuis and Friele (2005). A total of six creeks in the Flathead and Yahk watersheds were chosen based on biogeographic reasoning, access considerations and time/budget constraints (Friele and Dupuis 2006a).

Based on the results of the pilot it was decided that extensive monitoring should be limited to characterization of substrate texture and tadpole abundance. Due to low animal abundance, and high variance of both grain size and animal numbers, it was recognized that more work was required to determine appropriate sample sizes for trend detection. Further, it was stressed that sampling should be conducted in late summer when flows were low, but before stream temperatures dropped in the early fall, and that animal sampling might be distributed over several hundred metres, so as to disperse the impact of destructive area constrained sampling.

In August 2007, a three-year monitoring study was initiated on the six creeks sampled in the pilot study. The objectives of this work are to collect baseline data and refine the sampling methodology. This report summarizes the 2007 effort, compares the data with that collected in 2005, and makes recommendations for further work.

2.0 Guiding Principles

2.1 Importance of sample methodology

Gyug (2005, pg. 2) suggests that “area constrained searches are the closest thing to an absolute census” that there is, but emphasized the “closest” part, by noting multiple sweeps often turn up more tadpoles. He claims, “Until you can go through and be satisfied that there are no more tadpoles to be counted, you aren’t done.” Otherwise, sampling error is introduced into the estimate.

This can easily arise during area constrained searches (ACS) if animals are periodically missed due to lack of a downstream fence, incomplete channel-simplification and/or rubble rousing, failure to conduct sweeps, and when seasonality is introduced into sampling design. Seasonality affects detectability; with high detectability in the late-summer, cool to warm, low flow period, and low detectability in the spring freshet and fall/winter, cold low flow periods.

Therefore, adherence to a standardized and effective ACS search methodology is critical to ensure an abundance estimate is as close as possible to an absolute census. Similarly with stone counts, operator error may introduce a large, un-quantified error into grain size statistics (Bunt and Abt 2001). Again, the adherence to a standard and effective linear point count method is crucial to reduce operator error. For a description of search methods see Dupuis and Friele (2005) and Friele and Dupuis (2006a).

2.2 Significance levels and confidence limits

In ecology, testing at 90% confidence is often used (Toft and Shea 1983; Toft 1991), with a precision of about 20% about the mean reasonable for monitoring purposes (Gyug 2005). These levels were sought for both animal and grain size sampling.

3.0 Methods

3.1 Field

In the pilot study it was recommended that sample size for both texture and animal abundance be increased from the 2005 effort, which consisted of 100 stone point counts and 10, 3-m belts, or 30% effort, respectively. To disperse sampling along a longer creek length, an additional 100-m channel length was sampled at each site. The additional 100-m sample was located immediately upstream or downstream of the original, whichever was most similar in channel characteristics (gradient, morphology). All, except Boyd, were located upstream of the original; the Boyd sample was located downstream of the original and downstream of the road. With two 100-m reaches on each creek a total of 12, 100-m reach lengths were sampled. Within each 100-m reach, a 200 stone point count and 10, 3-m belt area constrained searches were conducted. Greater sample effort was not feasible given the budget.

The stone counts consisted of 2 parallel 100-m counts. A 100-m nylon tape was extended along one side of the channel from downstream to upstream, and stones selected at 1-m increments using a pointed brass rod as a plumb bob. Selected stones were measured using a square holed template, or "gravelometer". Then the tape was moved over at least 1-m, and the process repeated for a total sample of 200 stones.

For animal sampling, the 3-m belts were selected randomly, but with no overlap allowed. Animal searches used the area constrained, channel simplification methodology described in Friele and Dupuis (2006a). A 1/8" wire mesh fence was erected at the bottom of each belt sample. Qualitative embeddedness estimates were made at the end of each belt sample. All sampling was conducted from August 20-30, 2007. In all creeks, water temperatures were above 8ºC.

3.*2 Office*

For summaries and statistical tests, 2005 data was compiled with 2007 data, to allow comparisons between creek, reach and year. For abundance estimates, summary statistics and significance tests are calculated for tadpoles only, because frogs are not confined to the creek, and their numbers can vary by weather conditions.

The variable tadpole abundance is highly skewed, with numerous zero counts, and cannot be transformed to a normal distribution via LOG or SQRT transforms. Chi square test suggests the use of the negative binomial both to estimate sample size and test significance. As a result, for each creek the non-parametric Mann-Whitney U test was used to test differences between years. The required sample size for 90% confidence limits 25% about the mean was not calculated.

Pebble count data were grouped into grain size categories based on the phi scale (i.e., $-Log_2$): <2 mm; 2-4 mm; 4-8 mm; 8-16 mm; 16-32 mm; 32-64 mm; 64-128 mm; 128-256 mm; 256-512 mm; 512-1024 mm; and 1024-2048 mm. For each 100-m reach, grain size data were plotted as histograms, and on cumulative frequency diagrams so that percentiles could be determined for calculation of summary statistics.

From the frequency distributions, the graphic mean (MZ) and standard deviation (SI) of the samples were derived as follows (after Folk and Ward 57):

$$MZ = (D_{16}+D_{50}+D_{84})/3$$

$$SI = (D_{16}-D_{84})/4+(D_5-D_{95})/6.6$$

where D_i is the grain size in phi units of the *i*th percentile on the cumulative frequency plot (Folk and Ward 1957).

The graphical standard deviation provides a measure of sediment sorting, with four classes of sorting based on graphical standard deviation: SI = 1-2, poorly sorted (PS); 2-3, very poorly sorted (VPS); 3-4, extremely poorly sorted (EPS); >4, unsorted (US).

Mean and standard deviation were then used to calculate the required sample size with 90% confidence intervals 0.3 phi about the mean, following methods outlined in Bunt and Abt (2001).

For tailed frog work, as for fish, we are really interested in the fine tail of the distribution. However, due to limitations of the sampling method and operator errors, which both make tests of the fine tail less reliable (Bunt and Abt 2001), shifts of the mean were used for analysis, with significant shifts toward the fine tail causing concern. Differences between means was calculated using the T-test as outlined by Sokal and Rolf (1987), using formulas for both equal and unequal sample sizes:

$$\text{Equal n: } Ts=(mean_1-mean_2)/((1/n)*(stdev_1+stdev_2))^{1/2}, \text{ or}$$

$$\text{unequal n: } Ts=(mean_1-mean_2)/((stdev_1/n_1+stdev_2/n_2))^{1/2}$$

Where mean and stdev are measured in phi units (Table 3), and Ts and Tc refer to the sample T and the critical T value, respectively.

4.0 Results and Discussion

4.1 Survey effort

In the field, work estimates are for a two person crew. Due to the remoteness of the study area, at least 2-3 days were allocated for travel to and from the site (depending on distance from home), and 1 day for travel between Flathead and Yahk. Flathead samples were conducted from Aug 20-25 (6 days), and Yahk samples from Aug 27-30 (4 days). This amounts to a total of 14 days in the field for the 2007 season. Based on an 8-hour day (not including travel between camp and site), this yields 80 hours for 12, 100-m reaches, or 6 1/2 hours each. Given that the 200 stone point count required about 1 hour, the animal sampling took 5 1/2 hours per 30-m aggregate, 30 minutes per 3-m belt, or 10 minutes per metre.

In the office work estimates are for one person. Data compilation took one day for grain size analyses and one day for animal abundance. Statistical analyses and report writing require and additional 2-3 days, for a total of five days data compilation and report preparation.

4.2 Reach 2 descriptions

The assumption in extending the total sample reach another 100 m, is that the two 100 m segments are similar, so that reach effects are absent. However, on several creeks, differences between 100-m reaches could not be avoided. For example, R2 in Storm had about 10% bedrock substrate, and R2 in Cabin, Couldrey and Norge supported 1-2 large coarse woody debris jams with fine gravelly sediment wedges upstream. These differences could cause strong reach effects in grain size data. Reach effects in abundance estimates will be indirect.

4.3 Abundance estimates

The raw data from area constrained searches is supplied in an excel spreadsheet (Appendix 1), and the summary statistics are provided in Table 1. Considering all creeks, the number of animals caught per 3-m belt ranged from 0-37 tadpoles and 0-5 frogs. Tadpole numbers are highly skewed, with numerous zero counts, and generally low abundance (Total sample, Fig. 1).

Lognormal or square root transformations do not normalize the raw data. Chi square tests yield high values (chi>300) and suggest an aggregated distribution appropriate for use of the negative binomial distribution. Due to the inapplicability of the normal distribution, no attempt was made to estimate sample size from mean and standard deviation.

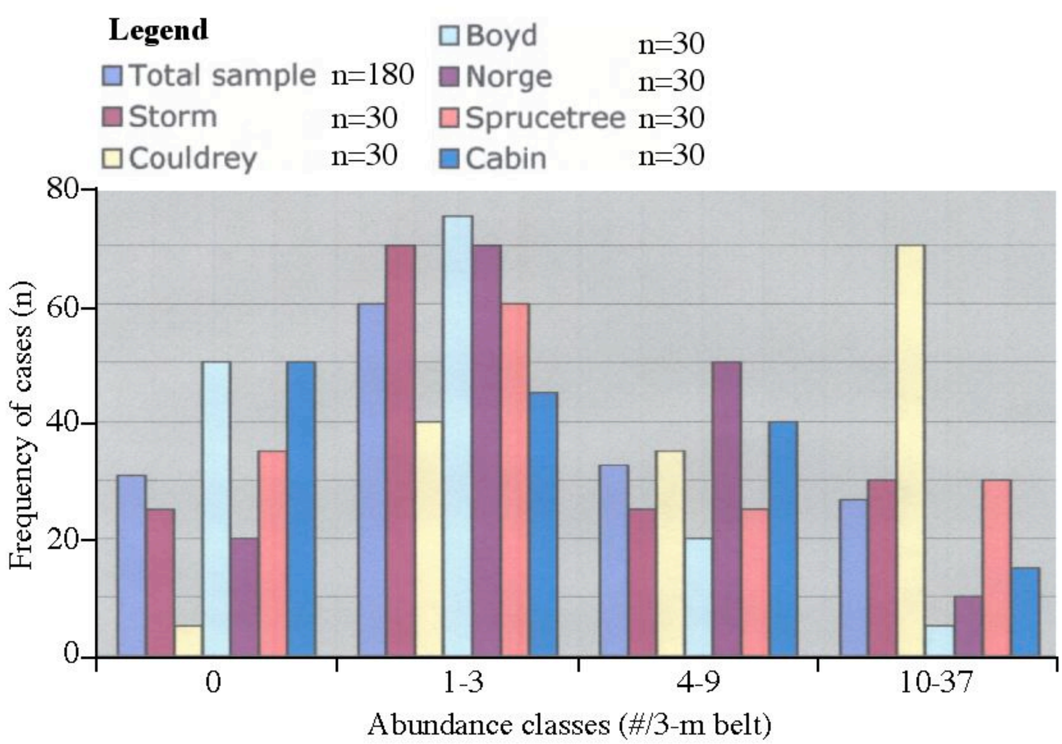

Figure 1. Tadpole abundance binned into frequency classes, with class boundaries subjectively selected to normalize frequency distribution. Total sample has been scaled (180/6) for comparison.

Binning abundance data into categories was conducted, with class boundaries subjectively selected to normalize the distribution (Fig. 1). The histogram for the total sample (n=180) remained left skewed, with zero counts making up 20% of the sample, and the mode in the 1-3 tadpoles/3-m belt category. The histograms by creek (n=30)(Fig. 1) indicate Storm and Sprucetree have similar distribution to the total sample, while Couldrey is skewed right with few zero data and many high counts, Boyd and Cabin are skewed left, with more zero data than the total, and Norge is more centrally distributed.

Table 1. Results of area constrained searches in 100-m sample reaches (n=10, 3-m belts). R1 is the 100-m reach sampled in 2005, R2 is the additional reach sampled in 2007. Grey highlighted rows are combined R1 and R2 in 2007 (n= 20, 3-m belts). Tad, tadpoles; frog, metamorphosed animals; TAFR, combined tadpoles and frogs.

Creek	Total tad	Total frog	Total TAFR	Mean tad	StDev tad	# of zero tad counts	Estimated # of tads/100 m	Mean embeddedness rank
Boyd/05-R1	10	7	17	1.0	1.2	4	33	3.2
Boyd/07-R1	37	3	40	3.7	3.3	1	123	3.5
Boyd/07-R2	8	0	8	0.8	0.9	5	27	3.5
Boyd/07-R1&2	45	3	48	2.3	2.8	6	75	3.5
Norge/05-R1	19	1	20	1.9	2.4	3	63	1.0
Norge/07-R1	69	12	81	6.9	5.7	1	230	1.2
Norge/07-R2	36	15	51	3.6	2.2	0	120	1.4
Norge/07-R1&2	105	27	132	5.3	4.5	1	175	1.3
Sprucetree/05-R1	15	1	16	1.5	1.9	4	50	1.1
Sprucetree/07-R1	85	8	93	8.5	5.4	0	283	2.1
Sprucetree/07-R2	55	10	65	5.5	7.8	3	183	2.1
Sprucetree/07-R1&2	140	18	158	7.0	6.7	3	233	2.1
Cabin/05-R1	14	1	15	1.4	2.1	4	47	2.7
Cabin/07-R1	35	5	40	3.5	3.4	3	117	2.6
Cabin/07-R2	47	3	50	4.7	6.0	3	157	3.3
Cabin/07-R1&2	82	8	90	4.1	4.8	6	137	2.9
Couldrey/05-R1	24	2	26	2.4	2.0	1	80	1.3
Couldrey/07-R1	96	0	96	9.6	4.5	0	320	2.8
Couldrey/07-R2	168	0	168	16.8	9.0	0	559	2.4
Couldrey/07-R1&2	264	0	264	13.2	7.8	0	440	2.6
Storm/05-R1	115	14	129	11.5	12.0	0	383	2.7
Storm/07-R1	47	0	47	4.7	7.4	4	157	3.0
Storm/07-R2	32	2	34	3.2	2.3	1	107	3.0
Storm/07-R1&2	79	2	81	4.0	5.4	5	132	3.0

The non-parametric Mann-Whitney U test, which makes no assumptions about normality, was used to test for differences in tadpole abundance between years. Data are tested by creek, indicate all creeks except Storm and Couldrey yielded higher abundance in 2007 (Table 2).

Table 2. Mann-Whitney U test for year effect (n=10)

Creek	M-W U	Wil W	Z	Sig.
Storm	25.5	80.5	-1.874	0.063*
Couldrey	5.5	60.5	-3.384	0.000
Boyd	18.0	73.0	-2.467	0.015
Norge	19.0	74.0	-2.362	0.018
Sprucetree	9.0	64.0	-3.141	0.001
Cabin	33.0	88.0	-1.320	0.218

* not significant with Kolmogorov-Smirnov test.

The increased abundance between years 2005 and 2007 is possibly related to seasonality: water temperatures were cold (<8ºC) for 2005 sampling and cool/warm (9-16ºC) for 2007 sampling. In 2005, many tadpoles may have already retreated to search-inaccessible winter refuge sites. That Boyd is the least populous creek is consistent with the fact that it supports the highest embeddedness values (Table 1) and finest substrates (Table 3).

4.4 Grain size analyses

The raw data for grain size analyses are presented in an excel spreadsheet (Appendix 2), and the summary statistics are presented in Table 3 and results of t-tests between year and between reaches are presented in Table 4. On all creeks the mean grain size is in the refuge-forming category (32-256 mm), as defined by Friele and Dupuis (2006a). Boyd Creek supports the finest substrate; while, Norge, Sprucetree, Cabin and Couldrey creeks are coarsest. Storm Creek is intermediate between these two groups.

Table 3 indicates that a sample size of 200 is somewhat larger than required for 90% confidence limits of 25% about the mean. T-tests suggest a significant difference in mean grain size from 2005 to 2007 on Boyd and Couldrey creeks. In Boyd fining occurred, while Couldrey appeared more coarse. Reach 2 in Norge and Cabin creeks was finer than reach 1. The finer texture of reach 2 is likely related to the existence of sediment aggradations behind coarse woody debris jams on reach 2.

Table 3. Grain size distribution summary statistics, sorting, sample size estimates, and confidence limits.

Creek	Sample size (n)	Mean (mm)	Mean (phi)	StDev (phi)	Sorting	Estimated n (90%; 0.3 phi)	Confidence limits (90%; 0.3 phi) Absolute (mm)		% of mean	
Boyd-R105	100	63	-5.974	-2.295	VPS	161	51	77	19%	23%
Boyd-R107	200	42	-5.397	-2.419	VPS	176	34	52	19%	23%
Boyd-R207	200	47	-5.563	-2.627	VPS	207	38	58	19%	23%
Boyd-R1,2/07	400	43	-5.427	-2.322	VPS	162	35	53	19%	23%
Norge-R105	100	131	-7.031	-1.628	PS	81	106	161	19%	23%
Norge-R107	200	132	-7.040	-1.547	PS	72	107	162	19%	23%
Norge-R207	200	102	-6.667	-1.522	PS	70	83	125	19%	23%
Norge-R1,2/07	400	116	-6.861	-1.529	PS	70	94	143	19%	23%
Spructree-R105	100	84	-6.390	-1.601	PS	78	68	103	19%	23%
Spructree-R107	200	101	-6.662	-1.391	PS	58	82	125	19%	23%
Spructree-R207	200	109	-6.766	-1.603	PS	77	88	134	19%	23%
Sprucetree-R1,2/07	400	106	-6.732	-1.437	PS	62	86	131	19%	23%
Cabin-R105	100	90	-6.489	-1.279	PS	50	73	111	19%	23%
Cabin-R107	200	101	-6.655	-1.788	PS	96	82	124	19%	23%
Cabin-R207	200	45	-5.498	-1.946	PS	114	37	56	19%	23%
Cabin-R1,2/07	400	70	-6.119	-1.956	PS	115	56	86	19%	23%
Couldrey-R105	100	117	-6.873	-1.775	PS	96	95	144	19%	23%
Couldrey-R107	200	95	-6.565	-1.794	PS	97	77	117	19%	23%
Couldrey-R207	200	84	-6.386	-2.286	VPS	157	68	103	19%	23%
Couldrey-R1,2/07	400	90	-6.487	-1.995	PS	120	73	110	19%	23%
Storm-R105	100	72	-6.162	-1.759	PS	95	58	88	19%	23%
Storm-R107	100	58	-5.862	-2.378	VPS	173	47	72	19%	23%
Storm-R207	100	63	-5.972	-2.256	VPS	156	51	77	19%	23%
Storm-R1,2/07	200	62	-5.962	-2.124	VPS	136	51	77	19%	23%

Note: PS, poorly sorted; VPS, very poorly sorted.

Table 4. Results of T-tests on mean grain size for each creek, by year and by reach.

Between years: Reach 1 only (2005, n=100; 2007, n=200)

Creek	Ts	df	Tc(0.1/200)	Sig
Boyd	3.08	298	1.65	**
Norge	-0.06	298	1.65	
Sprucetree	-1.79	298	1.65	
Cabin	-1.13	298	1.65	
Couldrey	1.88	298	1.65	*
Storm	1.48	198	1.66	

Between reaches: 2007 only (n =200)

Creek	Ts	df	Tc(0.1/200)	Sig
Boyd	-1.05	398	1.65	
Norge	3.01	398	1.65	**
Sprucetree	-0.85	398	1.65	
Cabin	8.47	398	1.65	**
Couldrey	1.26	398	1.65	
Storm	-0.51	198	1.66	

Ts, sample T; Tc, critical value of T for 90% confidence and appropriate n.
*, significant; ** very significant.

5.0 Conclusions

5.1 Survey effort

An effective monitoring program must find a balance between time/cost investments and sample precision. It is generally regarded that 1 day per creek (8 hours actual sampling, plus travel) be the maximum effort required by monitoring. Therefore, the work required to execute 20, 3-m reaches per creek (13 hours plus travel) is too much effort. Assuming 10 minutes per metre, then 10, 5-m belts might take about 8.3 hours. With about an hour for a 200 stone point count, this is 9 hours, plus travel. This is a potential alternate sampling strategy.

In future, assuming 1 field day per creek (n=6), plus 3 days for travel return and between, and 5 days write up, the effort to execute single year field program is 14 days; or 18 person days in the field, 5 person days in the office, plus expenses (mileage, accommodation, food). This could be extended if continued experimentation is desired in the 3-year baseline period to refine and improve methods.

5.2 Detecting trends in animal abundance

The problem of non-normality and high variance arises because, in general, zero counts are frequent, animal density is low (<4 tadpoles/3-m belt), but range is high (0-40). High counts (n=10-30) are often clusters; for example, nest sites with hatchlings or first year tadpoles. This aggregated abundance requires negative binomial analyses, or non-parametric analyses.

The fact abundance was stable (or declined) between years 2005 and 2007 in Storm Creek and increased in all the others, suggests the occurrence of a disturbance on Storm Creek that did not affect the others. The August 2003 Ram-Cabin fire is the only known recent disturbance unique to Storm Creek. Resampling of the 2003 inventory sample sites in 2004 (Friele and Dupuis 2006b) did not detect any difference in tadpole numbers in the sample reach or downstream, but it was concluded that there was frog mortality, especially upstream in the severely burned headwaters. Therefore, the apparent decline in tadpoles in Storm Creek between 2005 and 2007 could reflect a delayed impact of the 2003 Ram-Cabin fire. This could arise from reduced recruitment resulting from a decline in the number of breeding frogs. This trend will be interesting to monitor.

5.3 Detecting trends in grain size

A sample size of 200 stones per 100-m reach is appropriate to test difference between sample means at 90% confidence. The extra sampling, above the estimated required sample size (Table 3), will help to account for operator sampling error, which is known to be an important but ill-defined additional error (Bunt and Abt 2001). Reviewing the range of estimated sample size (Table 3) indicates sample size of 200 is less critical for the coarse creeks, but is important to detect fining, as fining increases the variance.

For all creeks, the mean grain size falls in the refuge-forming category (Table 3), consistent with the assumption that tailed frogs are adapted to these types of streams. All means are coarser than 40 mm. Boyd has the finest bed texture, with a mean between 40-60 mm; for Storm and Cabin R2 the mean is between 60-85 mm; while for the others it is between 85-130 mm. Based on spread of sample means across creeks, and known tadpole abundance responses to fines, some monitoring response rules may be defined:

- i) if mean grain size shifts from fine to coarse (Couldrey 05-07) then no concern is warranted;
- ii) if any significant fining is noted between years, a probable cause should be sought (Boyd),
- iii) if there is a shift of mean grain size from >60 mm to 40-60 mm, then moderate concern is warranted (Boyd); and
- iv) if mean grain size becomes <40 mm, then the creek bed is dominated by refuge-filling substrate, and serious concern is warranted.

With difference between reaches on Norge and Cabin, it is not valid to group reach 1 and reach 2 samples into a 400 stone point count.

6.0 Recommendations

1) Monitoring should be based on resampling of the originally defined 100-m reaches. Adding an additional 100-m reach to increase sample size to 20, 3-m belts and 400 stones increases the required effort above the “1-day per creek” goal.
2) Consultation with a biostatistician is recommended to search for alternate ways to estimate precision and test for significance of abundance data.

7.0 References

Bunte, K. and Abt, S.R.. 2001. Sampling surface and subsurface particle-size distributions in wadable gravel- and cobble-bed streams for analyses in sediment transport, hydraulics and streambed monitoring. United States Department of Agriculture. Rocky Mountain Research Station, General Technical Report RMRS-GTR-74. 428 p.

Dupuis, L.A., and Friele, P.A., 2005. Rocky Mountain Tailed Frog Monitoring Protocol. Contract No. CBIO5089. Project No. 2981052. Report to Kathy Paige, Ministry of Water, Land and Air Protection, Biodiversity Branch, Victoria, B.C.

Folk, R.L., and Ward, W.C., 1957. Brazos River bar: a study in the significance of grain size parameters. Journal of Sedimentary Petrology 27: 3-26.

Friele, P.A. and Dupuis, L.A., 2006a. Rocky Mountain Tailed Frog Monitoring: Results and Conclusions of the 2005 Pilot Study. MoF Contract CNLWL06043 Report to Kathy Paige, Biodiversity Branch, Environmental Stewardship Division, Ministry of Water, Land and Air Protection, Victoria, BC.

Friele, P.A. and Dupuis, L.A., 2006b. Post-fire assessment of tailed frog abundance in Storm and Leslie creeks, following the 2003 Ram-Cabin Fire. Report to Kathy Paige, Ministry of Water, Land and Air Protection, Biodiversity Branch, Victoria, B.C.

Gyug, L. 2005. Tailed frog tadpole population estimation methods, Wildlife Habitat Area monitoring pilot project. For Ecosystems Branch, BC Ministry of Environment, Victoria, BC.

Krebs, C.J. 1989. Ecological Methodology.Harper and Row, New York. 654 p.

Sokal, R., and Rolf, F.J., 1987. Introduction to Biostatistics, second edition. W.H. Freeman and Company, New York.

Toft, C.A. 1991. Reply to Seaman and Jaeger: an appeal to common sense. *In* Points of view: a controversy in statistical ecology. Herpetologica 46(3): 357-361.

Toft, C.A. and P.J. Shea. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. Amer. Nat. 122(5): 618-625.