Reader:

Please note the following errata for this article.

Erratum 1: page 759, right column, line 10. The equation within the text should read

$$R_i = N_i \, / \, N_{i-3}.$$

Erratum 2: page 760. Equation 1 should read

$$t = \frac{\bar{r} - \rho_{goal}}{s \, / \, \sqrt{n}}.$$

Erratum 3: page 761. Equation 2 should read

$$\delta = \frac{\rho - \rho_{goal}}{\sigma \, / \, \sqrt{n}}.$$

Abstract—When monitoring endangered species, natural resource managers require a recovery benchmark and a statistical procedure to test whether the benchmark has been met. We applied statistical power analysis to devise such a procedure for the endangered Sacramento River winter chinook salmon (*Oncorhynchus tshawytscha*). Winter chinook salmon management currently focuses on population growth rate, and our procedure used a Student's *t*-test to evaluate whether the average population growth rate is significantly lower than the management goal of 0.57 per generation. In the test, the null hypothesis was that the growth rate was not lower than the desired rate. In contrast to the usual hypothesis-testing framework, our procedure did not control for the type-I error rate. Instead, it controlled for the statistical power (the complement of the type-II error rate) and uses the resulting type-I error rate, computed from the sample size and other information, for the test. This procedure is conservative for winter chinook salmon in that, if all assumptions are met, it provides the specified level of assurance of detecting dangerously low population growth rates.

Monitoring protocol for Sacramento River winter chinook salmon, *Oncorhynchus tshawytscha*: application of statistical power analysis to recovery of an endangered species*

Steven T. Lindley

Michael S. Mohr

Michael H. Prager

Santa Cruz/Tiburon Laboratory
Southwest Fisheries Science Center
National Marine Fisheries Service, NOAA
3150 Paradise Drive
Tiburon, California 94920
E-mail address (for S. T. Lindley): Steve.Lindley@noaa.gov

The Sacramento River winter chinook salmon is listed as an endangered species under the U.S. Endangered Species Act (ESA). The historical spawning grounds of the winter chinook salmon were in upper tributaries of the Sacramento River, including the Upper Sacramento, Pit, and McCloud Rivers (Fig. 1). The completion of Shasta and Keswick Dams in the 1940s blocked access to these spawning grounds, although populations had already declined from historic levels owing to habitat destruction in the upper tributaries (Fisher, 1994). Quantitative winter chinook salmon population size estimates began in 1967 when the Red Bluff Diversion Dam (RBDD), a flashboard dam with three fish ladders, was completed. Since the completion of RBDD, winter chinook salmon spawning runs have declined from over 100,000 adults to a few hundred adults in the 1980s (Fig. 2; Williams and Williams, 1991).

The winter chinook salmon population remains extremely depleted. The California Fish and Game Commission listed the population as a "candidate" species under California's Endangered Species Act in 1988 and declared it endangered under that Act in 1989. The National Marine Fisheries Service (NMFS) declared the species "threatened" under the federal ESA in the same year, and it was declared "endangered" in 1994. NMFS has taken numerous regulatory actions under the ESA to improve winter chinook salmon survival, including changes in the regulations governing California's ocean salmon fisheries.[1]

In 1997, NMFS required that future ocean fishery harvest regulations be designed to achieve at least a 31% increase in the winter chinook salmon average cohort replacement rate over that observed in 1989–93.[2] Because winter chinook salmon females spawn predominantly at age 3 (Fisher, 1994), the cohort replacement rate in year i is, for simplicity, defined as $R_i = N_i/N_{i-3}$, where N_i is the number of adult spawners passing RBDD in year i. For statistical modeling purposes, it is convenient to express this cohort replacement rate on the log scale, $r_i = \log(R_i)$, and refer

---

* Contribution 103 of the Santa Cruz/Tiburon laboratory, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 3150 Paradise Drive, Tiburon, CA 94920.

1 Pacific Fishery Management Council. 1998. Review of 1997 ocean salmon fisheries. [Available from Pacific Fishery Management Council, 2130 SW Fifth Avenue, Suite 224, Portland, OR 97201.]

2 NMFS, 1996. Endangered Species Act-Section 7: Consultation. Biological opinion. The fishery management plan for commercial and recreational salmon fisheries off the coasts of Washington, Oregon, and California of the Pacific Fishery Management Council. National Marine Fisheries Service, Northwest and Southwest Regional Fishery Management Divisions. NMFS, 1997. Reinitiated Section 7: Consultation on the fishery management plan for commercial and recreational salmon fisheries off the coasts of Washington, Oregon, and California as it affects the Sacramento River winter chinook salmon. National Marine Fisheries Service, Northwest and Southwest Regional Fishery Management Divisions.

to r_i simply as the "growth rate" of the cohort returning in year i. Denoting the underlying mean growth rate by ρ, the goal of at least a 31% increase in the average cohort replacement rate over the observed mean of 1.35 for the 1989–1993 period is equivalent to a goal, on the log scale, of $\rho_{goal} \geq \log(1.35 \times 1.31) = 0.57$.

A natural recovery benchmark then is to compare the observed sample mean growth rate $\bar{r}$ in future years to ρ_{goal}. If recovery efforts have the desired effect, and if no increased mortality occurs from other causes, recovery will proceed as desired and $\bar{r}$ will likely exceed ρ_{goal}. However, if $\bar{r} < \rho_{goal}$, recovery may not be proceeding as desired and further conservation measures may need to be implemented. This possibility raises the question of how one should evaluate whether an observed $\bar{r} < \rho_{goal}$ warrants concern.

In our paper, we propose using a one-sample, one-sided t-test to evaluate the statistical significance of a difference between the observed mean growth rate and the target rate. We depart from the usual procedure (Lehmann, 1986), however, by conditioning the test on a specific level of statistical power, rather than on a fixed type-I error rate, in order to provide an adequate detection probability for dangerously low population growth rates. Application of the test requires choosing a particular power level and specifying what constitutes a "dangerously" low population growth rate. Together, these quantities determine the sensitivity of the test for detecting population growth rates below the target, and the likelihood of false positives, i.e. concluding that population growth rate is below the target when it in fact is not. The level of danger posed by a certain growth rate is evaluated by using a population viability model.

Figure 1

Map of Northern California depicting the Sacramento River, its upper tributaries, and mainstem dams.

Hypothesis testing and statistical power

Evaluating whether or not winter chinook salmon are meeting the recovery goal requires a statistical test because of the variability inherent in $\bar{r}$. We propose that a one-sided t-test be used to decide whether an observed mean population growth rate falls significantly short of the goal—in which case further regulatory action may be necessary. The null hypothesis of the test is that the underlying mean growth rate ρ (estimated by $\bar{r}$) is greater than or equal to ρ_{goal}, and the alternative hypothesis is that ρ is less than ρ_{goal}. That is,

$$H_0: \rho \geq \rho_{goal}$$

$$H_A: \rho < \rho_{goal}$$

with, in this case, $\rho_{goal} = 0.57$. Given a set of $n > 1$ observed r_i values $\{r_1, r_2, \ldots, r_n\}$, with mean $\bar{r}$ and standard deviation s, the test statistic is

$$\delta = \frac{\rho - \rho_{goal}}{\sigma / \sqrt{n}}. \tag{1}$$

Assuming that the $\{r_i\}$ are independent and identically distributed normal random variables, t has a central t-distribution with $n − 1$ degrees of freedom if $\rho = \rho_{goal}$, and a noncentral t-distribution if $\rho \neq \rho_{goal}$ (Lehmann, 1986). The t-test rejects H_0 in favor of H_A when t is less than some critical value t_c specified $a\ priori$ by the investigator.

A t-test has four possible outcomes, two of which result in correct inference: the test can accept H_0 when it is true, and it can reject H_0 when it is false. The test could also reject H_0 when it is in fact true—a "type-I" error, or it could fail to reject H_0 when it is false—a "type-II" error. The expected rates of type-I and type-II errors are conventionally denoted as α and β, respectively. The probability of correctly rejecting H_0 is known as the power π of the test, and $\pi = 1 − \beta$.

The type-I and type-II error rates are determined by the value of t_c selected for the test. A smaller value of t_c results in a lower type-I error rate α and a higher type-II error rate β. A larger value of t_c has the opposite effects. In all cases, the two error rates change in opposite directions when the value of t_c is changed. Thus, for a given data set of size n, α and β cannot be simultaneously minimized by adjusting t_c.

Given the endangered status of winter chinook salmon, it is clearly necessary to ensure that the statistical test has enough power to detect dangerously low population growth rates. To achieve this goal we propose that the power π of the test be held at a fixed level, rather than the

Figure 2
Historical time series of winter chinook salmon adult returns at Red Bluff Diversion Dam.
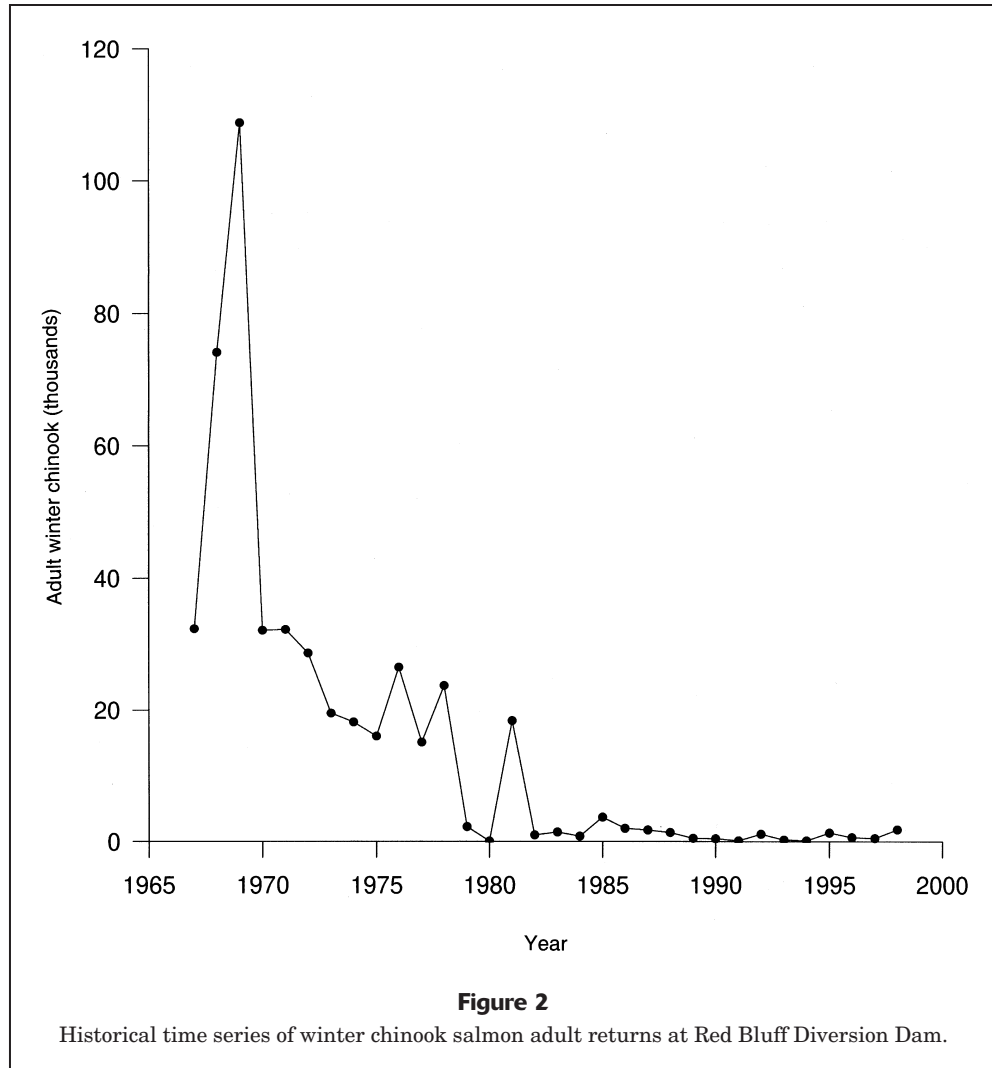
type-I error rate α. This approach differs from the conventional approach to hypothesis testing in which α is fixed at, say, 0.05, and the resulting power-level is either tolerated, or the sample size n is increased sufficiently to provide an acceptable level of power. Increasing n is not an option in this instance, because in any given year of application the sample size of the $\{r_i\}$ data set will be fixed, and a procedure is required that can be applied in each and every year. In the context of monitoring winter chinook salmon, a failure to reject H_0 may be used to justify "business as usual." By specifying the power π of the test in advance, resource managers will know that if the mean growth rate is falling seriously short of the goal, they will be able to detect this with specified probability π.

Calculation of α given π

In this section, we formulate the relation between π and α. This relation allows one to determine what values of α should be used for the test in order to achieve a specified level of power π.

With the previously stated distributional assumptions, t has a central t-distribution if $\rho = \rho_{goal}$, with cumulative distribution function (cdf) T and inverse cdf T^{-1}. If $\rho \neq \rho_{goal}$, t has a noncentral t-distribution with cdf T_δ, inverse cdf T_δ^{-1}, and noncentrality parameter

$$t = \frac{\bar{r} - \rho_{goal}}{s / \sqrt{n}}. \qquad (2)$$

which is the difference between ρ and ρ_{goal} in standard error units (Johnson et al., 1994). Given a particular critical value t_c, the associated type-I error rate and power of the t-test are

$$\alpha = \Pr\{\text{reject } H_0 | H_0 \text{true}\} = T(t_c), \qquad (3)$$

$$\pi = \Pr\{\text{reject } H_0 | H_0 \text{false}\} = T_\delta(t_c), \qquad (4)$$

α by definition being the largest value of $T_\delta(t_c)$ under H_0, which occurs at $\rho = \rho_{goal}$ where $T_{\delta=0}(t_c) = T(t_c)$. Solving

Equation 4 for t_c and substituting this into Equation 3 gives

$$\alpha = T\left(T_\delta^{-1}(\pi)\right). \qquad (5)$$

Equation 5 indicates the type-I error rate, α, associated with the test when conducted at specified power-level π. Note that, π having been specified, α also depends on the magnitude of the underlying difference $\rho - \rho_{goal}$, the variability σ in growth rates, and the sample size n, all through the noncentrality parameter δ.

Relation between α, π, n, and quasi-extinction

Although Equation 5 gives the value of α corresponding to a specified power-level π, the formula itself does not reveal the nature of the relation between α and π, and how this relation is affected by ρ and n. We illustrate these relationships below and consider their consequences in the context of a proposed test for detecting low growth rates in the Sacramento River winter chinook salmon population.

To apply Equation 5, we must specify σ, ρ, and π. We want to know (with probability π) that winter chinook salmon growth is not less than ρ_{goal} by a critical amount. Because the ESA is invoked to prevent extinction, we want to guard against growth rates that could lead to extinction. We used the winter chinook salmon population viability model developed by Botsford and Britt-nacher (1998) to identify growth rates corresponding to quasi-extinction probabilities of 0.05, 0.50, and 0.99 over 50 years. Quasi-extinction occurs when a population falls below some threshold level, in this case 200 adults in three consecutive cohorts. We initialized the viability model simulation with winter chinook salmon spawning escapements from the 1989–93 base period, and set $\sigma = 0.552$, the observed standard deviation of growth rates during the base period, assuming that this value will continue to hold in the future. The mean growth rates corresponding to the specified quasi-extinction proba-bilities were found to be about 0.0, –0.14, and –0.40, respectively. We note that if indeed $\rho = \rho_{goal} = 0.57$, quasi-extinction is an extremely unlikely event according to this model.

For each of the three growth rates, δ was computed according to Equation 2, and Equation 5 was then used to determine the type-I error rate α over a range of spec-ified power-levels π and sample sizes n. We calculated Equation 5 by using the Matlab Statistics Toolbox func-tion NCTINV (Jones, 1996). (Fortran and S-PLUS subrou-tines are also available for the noncentral t-distribution from the Carnegie Mellon University Department of Sta-tistics' StatLib and Oxford University Department of Sta-tistics' FTP archive, respectively. Alternatively, because t_δ is distributed as a ratio of independent random variables, $(Z + \delta) / \sqrt{\chi_{n-1}^2 / (n-1)}$, where Z is a standard normal variate and χ_{n-1}^2 is a chi-square variate with $n - 1$ degrees of free-dom, a large number of draws of t_δ could be simulated, and the $100 \times \pi$'th percentile could be taken as an approxima-tion to $T_\delta^{-1}(\pi)$ in Equation 5.)

Table 1

Type-I error rate (α) necessary to maintain the t-test power-level at π given a fixed number of observations n. Power-level is with respect to detecting a mean growth rate of ρ.

| ρ | π | n | | | |
		2	3	4	5
0.0	0.9	0.435	0.326	0.246	0.185
	0.8	0.308	0.216	0.154	0.110
	0.7	0.238	0.157	0.106	0.073
	0.6	0.188	0.117	0.076	0.050
	0.5	0.148	0.087	0.054	0.034
	0.4	0.114	0.063	0.037	0.023
	0.2	0.054	0.027	0.015	0.008
–0.14	0.9	0.334	0.225	0.152	0.102
	0.8	0.243	0.150	0.094	0.059
	0.7	0.190	0.109	0.065	0.039
	0.6	0.151	0.081	0.046	0.027
	0.5	0.120	0.061	0.033	0.018
	0.4	0.092	0.044	0.023	0.012
	0.2	0.044	0.019	0.009	0.004
–0.40	0.9	0.229	0.125	0.068	0.037
	0.8	0.172	0.084	0.042	0.022
	0.7	0.137	0.062	0.029	0.014
	0.6	0.110	0.046	0.021	0.010
	0.5	0.087	0.035	0.015	0.007
	0.4	0.068	0.025	0.010	0.004
	0.2	0.033	0.011	0.004	0.002

The results for $\rho = 0.0$ (Fig. 3) display the general behav-ior expected: 1) for fixed π, α decreases with n; 2) for fixed n, the higher π is set, the greater α becomes; and 3) for fixed α, power increases with increased sample size. For example, if the power-level were fixed at $\pi = 0.8$ and α set accordingly, a type-I error would be expected about 22% of the time with 3 years of data, 11% of the time with 5 years of data, or 3% of the time with 10 years of data. On the other hand, if the type-I error rate were fixed at $\alpha = 0.05$, there would be roughly an 80% chance of detecting this value of ρ with 7 years of data, but the power of detection would drop to $\pi < 35\%$ with only 3 years of data.

Table 1 lists results for mean growth rates of 0.0, –0.14 and –0.40. Notice that while the type-I error rate required diminishes for a given power-level and sample size as the underlying growth rate declines, use of a fixed $\alpha = 0.05$ even in the most dire case of $\rho = -0.40$ would provide very low power for $n \leq 3$.

Monitoring protocol

A monitoring protocol cannot be designed solely on the basis of statistical considerations—it must be guided by manage-ment policy. In this instance, the management policy is to
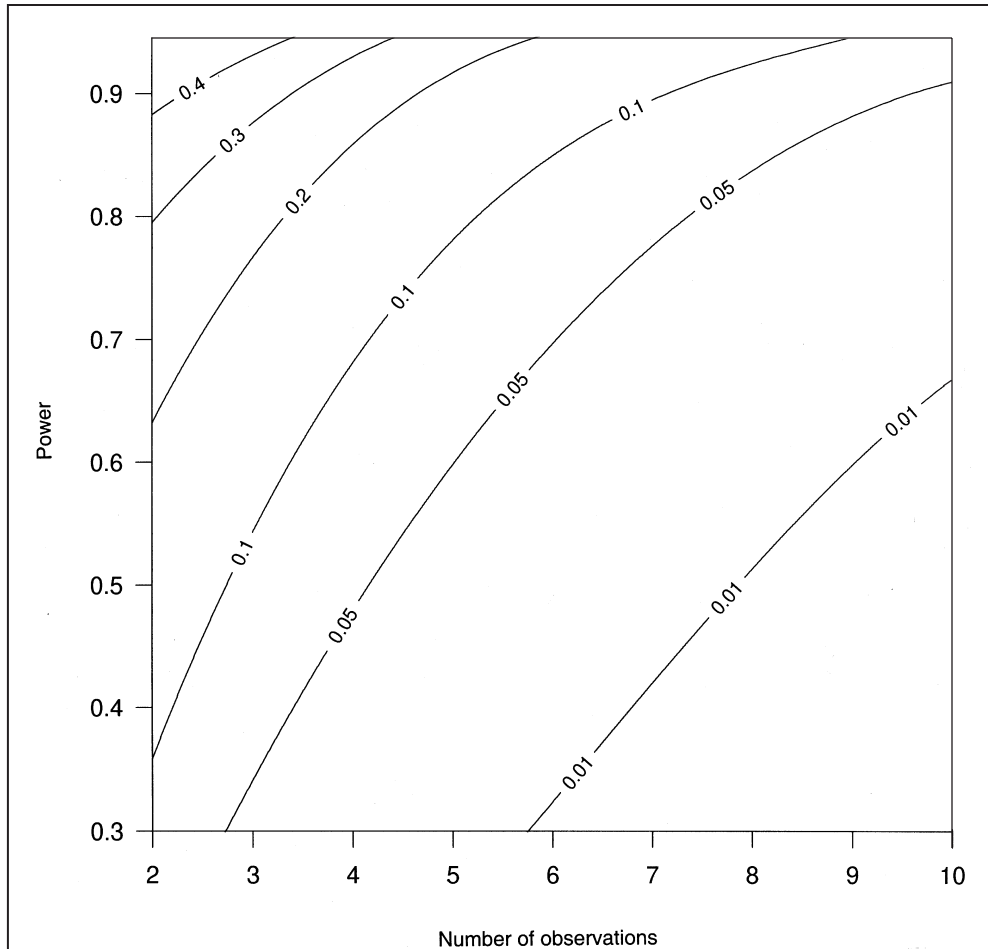
Figure 3

Type-I error rate (α) of the proposed *t*-test as a function of the power π and the number of observations n. Power-level is with respect to detecting a mean growth rate of $\rho = 0$ (which is projected to lead to quasi-extinction over 50 years with probability 0.05). The value of the noncentrality parameter in this case is $\delta = -0.57 / \left(0.552 / \sqrt{n} \right)$.

provide an adequate level of protection for winter chinook salmon through timely identification of low growth rates, without incurring too many false positive results.

The suggested protocol, therefore, would be to apply the *t*-test annually, with a fixed power-level of 80% for detecting a mean population growth rate of $\rho = 0$ (which is projected to lead to quasi-extinction over 50 years with probability 0.05). The choice of π and ρ is somewhat arbitrary, reflecting the perceived costs of type-I and type-II errors, and is discussed in a later section of this paper.

The observed growth rates are defined as $r_i = \log(N_i / N_{i-3})$, and the protocol would commence with the $i = 1997$ and 1998 spawning runs—both runs having benefited from the 1996 shift in ocean harvest regulations designed to reduce fishing mortality on winter chinook salmon.[1] Each year after 1998, the additional observed growth rates would be added to the test data set, until five growth rates are obtained. Beyond the year 2001, the test would be limited to the most recent five growth rates,

at which point the α-level will stabilize at 0.11. The protocol's moving five-year data frame will facilitate identification of shifts in winter chinook salmon survival and strengthen the basis for the assumption that the $\{r_i\}$ are identically distributed (discussed below). Survival shifts might be expected in response to naturally arising or management-related changes in the freshwater or marine environment.

To illustrate in concrete terms the proposed monitoring and analysis protocol, we applied it to the historical time series of adult returns $i = 1970, \ldots, 1996$, as if the protocol had commenced with the $i = 1970$ and 1971 spawning runs. The calculations and results of this application are presented in Table 2. Throughout the historical time series, $\bar{r}$ failed to reach the target level of 0.57, and this failure would have been declared significant in all years except for 1983–89 and 1996. If the test used a higher π, the number of failures declared significant would be higher. For instance, using $\pi = 0.85$ results in 21 null hypothesis

Table 2

Application of the proposed monitoring protocol to the historical time series of winter chinook salmon adult returns at Red Bluff Diversion Dam. Subscript i denotes year; N = abundance; R = cohort replacement rate; r = log (R) = growth rate; $\bar{r}$ = average growth rate over previous n years; t = t-statistic; π = power of test to detect a mean growth rate of zero; α = type-I error rate of test; t_c = test critical value; H_0 = mean growth rate ≥ 0.57.

Year(i)	N_i	N_{i-3}	R_i	r_i	$\bar{r}_i$	n_i	t_i	π_i	α_i	$t_{c,i}$	Reject $H_{0,i}$?
1967	32321	—	—	—	—	—	—	—	—	—	—
1968	74115	—	—	—	—	—	—	—	—	—	—
1969	108855	—	—	—	—	—	—	—	—	—	—
1970	32085	32321	0.993	−0.007	—	—	—	—	—	—	—
1971	32225	74115	0.435	−0.833	−0.420	2	−2.399	0.8	0.308	−0.689	yes
1972	28592	108855	0.263	−1.337	−0.726	3	−3.343	0.8	0.216	−0.976	yes
1973	19456	32085	0.606	−0.500	−0.669	4	−4.430	0.8	0.154	−1.225	yes
1974	18109	32225	0.562	−0.576	−0.651	5	−5.613	0.8	0.110	−1.453	yes
1975	15932	28592	0.557	−0.585	−0.766	5	−8.718	0.8	0.110	−1.453	yes
1976	26462	19456	1.360	0.308	−0.538	5	−4.252	0.8	0.110	−1.453	yes
1977	15028	18109	0.830	−0.186	−0.308	5	−5.160	0.8	0.110	−1.453	yes
1978	23669	15932	1.486	0.396	−0.129	5	−3.337	0.8	0.110	−1.453	yes
1979	2251	26462	0.085	−2.464	−0.506	5	−2.068	0.8	0.110	−1.453	yes
1980	84	15028	0.006	−5.187	−1.427	5	−1.859	0.8	0.110	−1.453	yes
1981	18297	23669	0.773	−0.257	−1.540	5	−2.041	0.8	0.110	−1.453	yes
1982	972	2251	0.432	−0.840	−1.671	5	−2.243	0.8	0.110	−1.453	yes
1983	1439	84	17.131	2.841	−1.182	5	−1.327	0.8	0.110	−1.453	no
1984	794	18297	0.043	−3.137	−1.316	5	−1.388	0.8	0.110	−1.453	no
1985	3633	972	3.738	1.318	−0.015	5	−0.579	0.8	0.110	−1.453	no
1986	2013	1439	1.399	0.336	0.104	5	−0.462	0.8	0.110	−1.453	no
1987	1761	794	2.218	0.797	0.431	5	−0.141	0.8	0.110	−1.453	no
1988	1386	3633	0.382	−0.964	−0.330	5	−1.129	0.8	0.110	−1.453	no
1989	480	2013	0.238	−1.434	0.011	5	−1.070	0.8	0.110	−1.453	no
1990	425	1761	0.241	−1.422	−0.537	5	−2.386	0.8	0.110	−1.453	yes
1991	134	1386	0.097	−2.336	−1.072	5	−3.173	0.8	0.110	−1.453	yes
1992	1122	480	2.338	0.849	−1.061	5	−3.096	0.8	0.110	−1.453	yes
1993	267	425	0.628	−0.465	−0.961	5	−2.832	0.8	0.110	−1.453	yes
1994	153	134	1.142	0.133	−0.648	5	−2.165	0.8	0.110	−1.453	yes
1995	1296	1122	1.155	0.144	−0.335	5	−1.670	0.8	0.110	−1.453	yes
1996	612	267	2.292	0.829	0.298	5	−1.101	0.8	0.110	−1.453	no

rejections out of 26 tests, rather than the 18 rejections for $\pi = 0.80$. On the other hand, setting $\alpha = 0.05$ would have rejected H_0 only 14 times.

Discussion

Statistical considerations

Any t-test assumes that the observations are independent and are identically distributed normal random variables. Assuming a normal distribution for r is reasonable, because R (its antilog) has been found to be approximately lognormally distributed (Botsford and Brittnacher, 1998). More generally, observations on fish population sizes or survival rates are often found to be approximately lognor-

mally distributed, as if arising from a series of random multiplicative events (Hennemuth, 1980).

The requirement that the growth rate observations be identically distributed as well (same underlying mean and variance), is not directly verifiable, but this assumption is reasonable given the short time period over which the test is conducted (n=5-year period). Indeed, this was our rationale for limiting the testing protocol to a 5-year time horizon. Although a longer time frame would boost the test's sample size and, for fixed α, increase its power, we believe identity of distribution beyond a 5-year horizon is an untenable assumption. Thus, the proposed protocol limits the test to a 5-year time frame, achieving the necessary power at the expense of a higher type-I error rate.

The issue of independence of observations is more difficult to assess. Winter chinook salmon adults return to

the river for spawning at age 3 or age 4, and although approximately 90% of a brood's spawning adults do so at age 3 (Fisher, 1994), this means that there is imperfect temporal isolation between runs in adjacent years. It is plausible that an environmental factor occurring in a particular spawning year could affect the returns in several future years, creating a lack of independence among observations. Such autocorrelation would decrease the effective sample size. Under this scenario, any test based on the nominal sample size (number of years of data) would have increased levels of both type-I and type-II errors (Lehmann, 1986). Given the relatively low level of information available on this population, this consideration is secondary and can be evaluated more thoroughly as the data base increases. The current time series of $\{r_i\}$ values shows no significant autocorrelation ($P>0.05$), which suggests that a lack of independence is most likely not a serious issue.

Choice of δ and π

In any hypothesis test, one must specify the type-I and type-II error rates. The choice of these rates should reflect the relative costs, as perceived by the investigator, of making these errors (Toft and Shea, 1983). When these costs can be specified in advance, and in comparable terms, one can balance them explicitly through specification of α and π (Mapstone, 1995). However, in the case of our proposed winter chinook salmon monitoring protocol, the cost of making a type-I error is unknown (no specific actions have yet been associated with a rejection of the null hypothesis), whereas type-II errors may be associated with extinction. We believe the appropriate course in this situation is to first identify growth rates that lead to unacceptably high probabilities of extinction, fix the power π of detecting these growth rate levels at a suitably high level, and accept, within reasonable limits, the resulting type-I error rate α. Specifically, for the winter chinook salmon monitoring protocol, we have specified an 80% chance of detecting growth rates that would lead to a ≥5% chance of quasi-extinction in 50 years, and accept the corresponding type-I error rate. We have selected these values because they are consistent with suggestions in the literature (reviewed briefly below); resource managers should carefully consider whether they are appropriate.

Setting δ by way of ρ that leads to an unacceptable predicted extinction risk, as we have done, is natural in the current setting, but just what level of extinction risk should be of concern is debatable. Population viability models have been widely used in conservation biology to quantify extinction risk as a function of population size and the magnitude and variability of population growth rate (Beissinger and Westphal [1998] have provided a recent review). Shaffer (1981), in pioneering work on minimum viable populations, has tentatively suggested that viable populations should have at least a 99% chance of remaining extant for 1000 years, but stated that specific probabilities and time horizons are arbitrary, and other values might be more appropriate. Indeed, other studies have used a variety of criteria: Botsford and Brittnacher (1998) used a 0.10 extinction probability in 50 years to develop criteria

for removing winter chinook salmon from the Endangered Species List; Shaffer and Samson (1985) used the criteria of a 0.05 extinction probability over 100 years to identify a minimum viable population size for grizzly bears. We have adopted the 0.05 probability of extinction over 50 years as a moderately conservative criterion.

Specifying the value of π is also somewhat arbitrary. Peterman and M'Gonigal (1992) contend that monitoring programs must have high power ($\pi \geq 0.8$) to detect biologically important effects in order to be reliable. A reliable test should also have a reasonable α value as well as sufficient power. In the proposed winter chinook salmon protocol, the α-level stabilizes at 0.11 after 5 years of data have been collected, and we feel that this behavior represents a reasonable balance between the type-I and type-II error rates.

Power analysis and the precautionary approach

With the decline, collapse, or endangerment of numerous fish populations around the world, the paradigm of precautionary fishery management is receiving increasing attention. The "precautionary approach" to fishery management, as developed by the Food and Agriculture Organization of the United Nations (FAO, 1996), strives to avoid irreversible or slowly reversible damage to fisheries, places priority on conservation of productive capacity, and requires that fishing activities be considered harmful unless proven otherwise. The reauthorization of the U. S. Magnuson-Stevens Fishery Conservation and Management Act (as amended through October 11, 1996) is meant to ensure that "irreversible or long-term adverse effects on fishery resources and the marine environment are avoided."

Peterman and M'Gonigal (1992) have argued that power analysis is a fundamental part of precautionary management, because it provides an estimate of the reliability of the monitoring program. There are four types of power analysis, which correspond to determining one of either n, π, α, or the effect size from the other three (Cohen, 1977). In environmental studies, the determination of n and π are fairly common (e.g. Gerrodette, 1987; Gryska et al., 1997; Urquhart et al., 1998). The determination of α, as we have done here, is least common, in part owing to the "strength of the significance criterion convention, which makes investigators loath to consider "large" values of α" (Cohen, 1977).

Type-II errors in fisheries management are costly because populations and ecosystems can be slow to recover (Dayton, 1998). In endangered species management, the biggest risk is extinction of a species, rather than failure to meet some fiscal or harvest goal, and is truly irreversible. Fixing the type-I error rate at a typical value such as 0.01 or 0.05 would make timely detection of dangerously low growth rates unlikely (Table 1; Peterman, 1990). Thus, we believe that using standard statistical protocols, which control for the type-I error rate and accept the resulting type-II error rate, is not an appropriate method when monitoring endangered species. In such situations, it is more logical, and certainly more precautionary, to set the type-II error rate at an acceptably small value that yields a reasonable type-I error rate.

Acknowledgments

Literature cited

Beissinger, S. R., and M. I. Westphal.
   1998.   On the use of demographic models of population viability analysis in endangered species management.   J. Wildl. Manage. 62:821–841.
Botsford, L. W., and J. G. Brittnacher.
   1998.   Viability of Sacramento River winter-run chinook salmon.   Conserv. Biol. 12:65–79.
Cohen, J.
   1977.   Statistical power analysis for the behavioral sciences. Academic Press, Inc. Orlando, FL, 474 p.
Dayton, P. K.
   1998.   Reversal of the burden of proof in fisheries management.   Science (Washington, D.C.) 279:821–822.
FAO.
   1996.   Precautionary approach to capture fisheries and species introductions. elaborated by the Technical Consultation on the Precautionary Approach to Capture Fisheries (Including Species Introductions) Lysekil, Sweden, 6–13 June 1995.   FAO Technical Guidelines for Responsible Fisheries 2, FAO, Rome, 54 p.
Fisher, F. W.
   1994.   Past and present status of Central Valley chinook salmon.   Conserv. Biol. 8:870–873.
Gerrodette, T.
   1987.   A power analysis for detecting trends.   Ecology 68: 1364–1372.
Gryska, A. D., W. A. Hubert, and K. G. Gerow.
   1997.   Use of power analysis in developing monitoring protocols for the endangered Kendall Warm Springs dace.   N. Am. J. Fish. Manage. 17:1005–1009.
Hennemuth, R. C.
   1980.   A statistical description of recruitment in eighteen selected fish stocks.   J. Northwest Atl. Fish. Sci. 1:101–111.
Johnson, N. L., S. Kotz, and N. Balakrishnan.
   1994.   Continuous univariate distributions, vol. 2, 2nd ed. John Wiley & Sons, Inc., New York, NY, 719 p.
Jones, B.
   1996.   Statistics toolbox user's guide.   The MathWorks, Inc., Natick, MA, 342 p.
Lehmann, E. L.
   1986.   Testing statistical hypotheses, 2nd ed.   John Wiley & Sons, Inc., New York, NY, 600 p.
Mapstone, B. D.
   1995.   Scalable decision rules for environmental impact studies: effect size, Type-I, and Type-II errors.   Ecol. Appl. 5:401–410.
Peterman, R. M.
   1990.   Statistical power analysis can improve fisheries research and management.   Can. J. Fish. Aquat. Sci. 47:2–15.
Peterman, R. M., and M. M'Gonigal.
   1992.   Statistical power analysis and the precautionary principle.   Mar. Pollut. Bull. 24:1231–1234.
Shaffer, M. L.
   1981.   Minimum population sizes for species conservation. Bioscience 31:131–134.
Shaffer, M. L., and F. B. Samson.
   1985.   Population size and extinction: a note on determining critical population size.   Am. Nat. 125:144–152.
Toft, C. A. and P. J. Shea.
   1983.   Detecting community-wide patterns: estimating power strengthens statistical inference.   Am. Nat. 122:618–625.
Urquhart, N. S., S. G. Paulsen, and D. P. Larsen.
   1998.   Monitoring for policy—relevant regional trends over time.   Ecol. Appl. 8:246–257.
Williams, J. E., and C. D. Williams.
   1991.   The Sacramento River winter chinook salmon.   Chapter 9 in California's salmon and steelhead: the struggle to restore an imperiled resource (A. Lufkin, ed.), p. 105–115. Univ. California Press, Berkeley, CA, 305 p.