DOI: 10.1007/s002670010202

PROFILE

Criteria for Oil Spill Recovery: A Case Study of the Intertidal Community of Prince William Sound, Alaska, Following the *Exxon Valdez* Oil Spill

JOHN R. SKALSKI*
School of Aquatic and Fishery Sciences
University of Washington, Box 358218
Seattle, Washington 98198-8218, USA

DOUGLAS A. COATS
Marine Research Specialists
3639 East Harbor Boulevard, Suite 208
Ventura, California 93001, USA

ALLAN K. FUKUYAMA
School of Aquatic and Fishery Sciences
University of Washington, Box 355020
Seattle, Washington 98195-5020, USA

ABSTRACT / Marine intertidal organisms in Prince William Sound were exposed to crude oil following the T/V *Exxon Valdez* oil spill in 1989. The intertidal communities were also subjected to mechanical disturbance during invasive oil spill remediation and cleanup efforts. Using monitoring data collected from 1989 to 1997, impacts and eventual recovery were assessed at oiled but uncleaned sites and oiled and cleaned study areas. A statistical model where recovery was defined as parallelism between the time profiles at control and oiled sites was evaluated. Statistical analysis and graphical presentations of the data suggest intertidal epibiota communities recovered from the oil spill by 1992 at the oiled sites and by 1994 at the oiled and remediated sites. Empirical data from the intertidal monitoring program supports the use of tests of parallelism in evaluating recovery and the need to avoid simply the comparison of sample means from control and oiled sites.

The primary considerations following a major oil or chemical spill are containment, cleanup, and impact assessment. However, shortly thereafter, attention naturally shifts to the long-term consequences of the spill event and eventual recovery. It is this evaluation of when and whether recovery has occurred at the intertidal communities in Prince William Sound following the *Exxon Valdez* oil spill that is the focus of this paper.

Holloway (1996) described how perspective and scale have influenced the scientific processes of assessing the recovery of the intertidal community in Prince William Sound following the *Exxon Valdez* oil spill. To some extent, recovery may be in the eye of the beholder, for recovery can be interpreted differently by different people. For example, Holloway (1996) interpreted Exxon's definition of recovery as "the reestablishment of a healthy biological community characteristic of the area," while summarizing for the state and federal resource trustees that "recovery will occur when the Sound looks as it would have if the spill had not occurred." Unfortunately, no detailed baseline information exists that can be used to accurately describe the sound prior to the spill, and no technology exists that can extrapolate those conditions a decade into the future. Hence, ten years after the *Exxon Valdez* oil spill, the debate continues over whether and where recovery has occurred.

The purpose of this paper is to describe alternative statistical end points that may be used to identify when an intertidal community has recovered from an event such as an oil spill. Nine years of postspill monitoring data collected by the National Oceanic and Atmospheric Administration (NOAA) in Prince William Sound will be used to illustrate the concepts. The design limitation of conducting an accident assessment (Skalski 1995), along with the biological processes of community dynamics, will help form our recommendations.

Study Area

In 1989, the *Exxon Valdez* accident spilled 41.6 million liters of oil in Prince William Sound and the outer coastal areas of the Gulf of Alaska. Approximately 35 million liters impinged on 645 km of shoreline and

KEY WORDS: Accident assessment; Impact; Intertidal; *Exxon Valdez*; Oil spill; Prince William Sound; Recovery

---

*Author to whom correspondence should be addressed; *email:* jrs@cbr.washington.edu became stranded on intertidal habitats (Spies and others 1996). Oil coated rock surfaces, penetrated into soft sedimentation, and potentially impacted a wide range of intertidal organisms. Shoreline cleaning removed a large amount of the stranded oil but also damaged the intertidal environment. High-pressure, hot-water washing of the rocky intertidal environment was particularly destructive (Mearns 1996).

Intertidal shoreline monitoring studies were implemented in 1989 shortly after the *Exxon Valdez* oil spill. Results of geomorphology and chemistry studies are provided by Michel and Hayes (1994), Hayes (1996), Roberts and others (1997), Hayes and Michel (1998). Results of the intertidal studies are available in several annual reports (e.g., Houghton and others 1997, Coats and others 1999). This paper presents some of the results of the field sampling from 1989 to 1997.

The primary objective of the NOAA intertidal studies was to examine the effects of oiling and cleanup on intertidal epibiota and infaunal communities and to document subsequent recovery processes. The majority of the study sites were selected in 1989 or 1990, based on their degree of oiling and cleanup history. Three categories of sites were selected for the investigation:

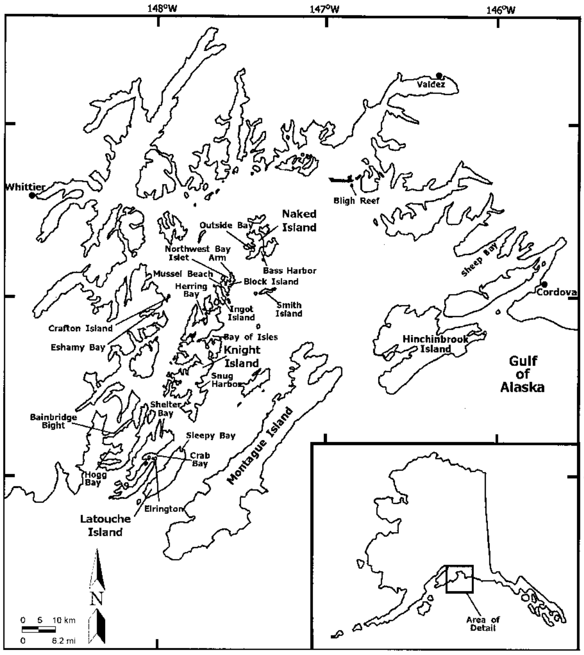

Figure 1. Intertidal sampling locations within Prince William Sound.

Category 1—unoiled with no cleaning treatments applied

Category 2—oiled and untreated or cleaned only with cool-water flushes in 1989

Category 3—oiled and treated with hot-water, high-pressure washes one to three times

The sampling sites were located primarily within the western portion of Prince William Sound, although one control site (Sheep Bay) was located to the east (Figure 1). The rationale for categorizing the stations is provided in Houghton and others (1993, Appendix A-1).

Methods

At epibiota sites, a stratified random sampling design was used to select sampling locations along transects in the upper, middle, and lower tidal levels. The transect lines were established according to biological communities at each tidal level. The upper transect was established within the *Verrucaria* (lichen) zone above where rockweed (*Fucus gardneri*) prevails. The middle transect was located within the *Fucus* zone, and the lower transect was positioned just below the *Fucus* zone near mean low water.

Epibiotic sampling was conducted within five (upper transect) or ten (middle transect) permanent 0.25-m^2 quadrats established along a transect. Epibiotic enumeration within each quadrate consisted of counts of individual organisms and estimates of species coverage. Percent cover was estimated for overstory, understory, and crustose algae, as well as mussels, barnacles, and encrusted invertebrates such as sponges and bryozoans.

Infauna were collected with a modified cylindrical clam corer of 10.5 cm diameter × 15 cm long. Replicate cores were collected at each of nine core stations along the lower intertidal transect only. Samples were live-sieved in the field using a 10-mm mesh size and preserved in 10% solution of buffered formalin. Upon arrival at the laboratory, samples were rinsed on a 0.5-mm mesh sieve, transferred to 70% ethanol and Rose Bengal, and sorted under a binocular dissecting microscope. Taxonomists identified organisms to lowest possible levels.

Definitions of Recovery

As is inherent to most accident assessment studies (Skalski 1995), no prespill baseline information was available for the intertidal communities exposed to spilled oil from the *Exxon Valdez*. A similar lack of information also characterizes most of the other communities exposed to oil from the *Exxon Valdez* accident (Wells and others 1995, Rice and others 1996). Hence, recovery cannot be simply based on the notion of a return to the intertidal population levels and composi-

Table 1. Summary of possible recovery end points proposed by the Exxon Valdez Oil Spill Restoration Plan (EVOSTC 1999)

Population or community being assessed	Recovery definition
Bald eagles, black oystercatcher, common loons, common murres, cormorants, harlequin ducks, sea otters	Population return to prespill levels.
Clams	Populations/productivity return to levels that would have prevailed in absence of oil spill, based on comparisons of oiled and unoiled sites.
Cutthroat trout, Dolly varden	Growth rates within oiled areas similar to those for unoiled areas, taking into account geographic differences.
Harbor seals, killer whales, marbled murrelets, pigeon guillemots	Population stable or increasing.
Intertidal communities	Community composition on oiled shorelines similar to that which would have prevailed in absence of the spill.
Subtidal communities	Community composition in oiled areas similar to unoiled areas.
Mussels	Concentrations of oil in mussels and in sediments below beds reach background, do not contaminate predators, and do not affect subsistence uses.
Pacific herring	The next highly successful year class is recruited into the fishery and when other indicators of population health are sustained within normal bounds.
Pink salmon	Population indicators (e.g., growth and survival) within normal bounds and no statistically significant differences in egg mortalities in oiled and unoiled streams for two years each of odd and even year runs.
Sockeye salmon	Adults returns per spawner and other indicators or [sic] productivity are within normal bounds.
River otters	Biochemical indices of hydrocarbon exposure or other stresses and indices of habitat use are similar between oiled and unoiled areas after taking into account any geographical differences.

tion prior to the spill, because the prespill levels are unknown. Nevertheless, the Exxon Valdez Oil Spill Trustees Council (EVOSTC 1999) "continues to use prespill numbers or conditions as the most useful benchmark in evaluating the status of recovery." Assessment also cannot be based simply on a comparison between mean response levels at control and impacted sites, because they may have been different prior to the spill. Instead, inferences concerning recovery must be based on indirect evidence that oil-exposed sites have returned to an unimpaired condition.

The lack of baseline information and absence of unbiased reference or control sites in assessing recovery has led to a proliferation of assessment strategies among environmental investigators. No less than 11 different definitions (Table 1) of recovery have been identified in the 1999 Exxon Valdez Oil Spill Restoration Plan (EVOSTC 1999). No unified strategy for assessing recovery appears to exist. Many of these recovery end points are dependent upon comparisons with prespill levels or based on simple comparisons between oiled and unoiled sites. A more reasonable approach, given the inherent limitations of accident assessments, is to define recovery as when the impacted populations or communities return to levels that would have prevailed in the absence of the oil spill. But how can we know when the prespill environment has returned? The next section describes an assessment strategy that is supported by nine years of postspill data from intertidal communities in Prince William Sound.

Parallelism Hypothesis

Under the conceptual model of an acute effect of an oil spill, followed by eventual recovery over time, time profiles of control and impacted sites would diverge upon impact and with sufficient time, begin to track or parallel each other as impacted sites begin responding solely to the same regional climatic changes or oceanographic conditions as the reference sites. Hence, eventual parallelism between mean profiles for control and oil-treatment sites would provide inferential evidence of recovery (Skalski and Robson 1992, pp. 194–211, Skalski 1995). From this perspective, recovery can be considered complete when impacted intertidal populations eventually begin to track or parallel the control site profiles. Under this scenario, the statistical tests of recovery are equivalent to tests for parallelism. Differences in population or community levels between control and impacted sites are not a consideration in assessing recovery, only the relative patterns of the temporal trends are of interest.

Focusing on the temporal trend is an important aspect of the recovery analysis because there is no a priori reason to expect mean population levels at control and impacted sites to be the same, even in the absence of the spill. On the contrary, arbitrary differences in average population levels are likely to occur because of natural variation, spatial confounding, and lack of randomization. Systematic environmental differences between control and impacted sites might be expected for reasons unrelated to oiling or shoreline cleanup techniques. These differences can arise from natural processes such as the prevailing current flow, as suggested by Spies and others (1996). Currents that direct oil to certain shorelines may also be responsible for the distribution of larvae and nutrients. Because these site differences cannot be randomized across treatment designations, recovery assessments based on the direct comparison of mean population levels are statistically untenable.

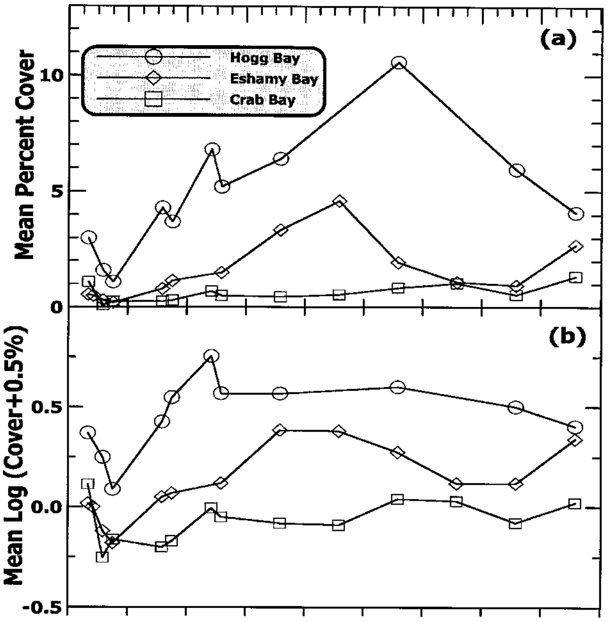

Figure 2. Time profiles for mean percent cover of the small acorn barnacle (*Chthamalus dalli*) at three control sites on the (**a**) arithmetic scale and (**b**) logarithmic scale.

Analysis on the Proper Scale

Prior to the statistical analysis, it is imperative that the field data be properly transformed to achieve additivity of the environmental effects (Bartlett 1947, Eberhardt 1978). The data need to be analyzed on the scale where natural differences between sites and temporal effects have an additive effect on population levels. Under these circumstances, the mean temporal profiles of the control and impacted sites will be parallel over time in the absence of effects from the spill. A test of recovery is then essentially a test of parallelism. Parallelism would be suggestive of recovery, while lack of parallelism would suggest the impacted sites have not yet recovered.

The temporal trends at the replicate control sites can be used to identify the proper scale of analysis. At control sites, the time profiles reflect localized conditions as well as responses to regional changes in oceanographic and climatic conditions that should affect all or most sites similarly in the Sound. If each of these control sites responds to the regional climatic changes over time similarly, then the temporal profiles should show similar trends over time as well. Actual population levels may differ among the control sites because of differences in local conditions, but the changes over time would be similar if these sites are responding to the same temporal changes in the regional environment. An important element of the proposed test of recovery is to determine the scale where the temporal profiles of the control sites are parallel.

For many species in Prince William Sound, a logarithmic transformation of the data suggests that control populations are tracking or paralleling each other on a multiplicative scale. For example, Figure 2 illustrates time profiles for percent cover of small acorn barnacles (*Chthamalus dalli*) at the three control sites. Profiles for mean percent cover (Figure 2a) suggest erratic changes in abundance over time across the three sites. However, after logarithmic transformation, the time profiles show differences in amplitude but also trends that more clearly parallel over the nine years of sampling. The time profiles for the abundance of the infaunal bivalve *Rochefortia tumida* also illustrate near parallelism after log transformation (Figure 3).

There are many reasons to compare population trends on a logarithmic rather than arithmetic scale. In particular, the processes of natality and mortality tend to have multiplicative effects on population abundance and density. Environmental conditions are also more likely to produce similar fractional changes in abundance rather than equal absolute changes in abundance. Geometric change has been a fundamental premise of population dynamics for 200 years (Malthus 1798). Toxic effects have also been implicitly acknowledged as having multiplicative effects on population abundance by the use of effective concentrations (e.g., EC_{50}) in toxicity tests.

Graphical examinations and Tukey's (1949) test of nonadditivity can be used to help select the proper transformation using the control site data. The control data can be analyzed by using site × year cross-classifi-

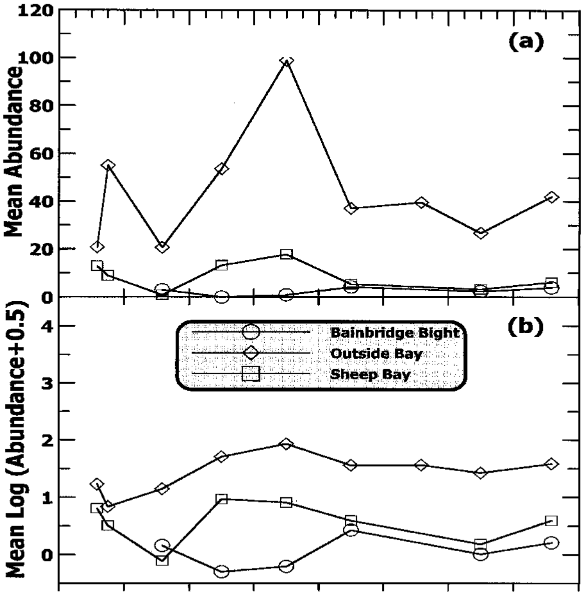

Figure 3. Time profiles for mean abundance of the bivalve *Rochefortia tumida* at the three control sites on the (**a**) arithmetic scale and (**b**) logarithmic scale.

cations in a two-way test of nonadditivity (Kirk 1982, pp. 250–253).

Statistical Methods

Returning to the same control and treatment sites year after year constitutes a multivariate repeated measures study (Morrison 1976, pp. 205–222). In the Prince William Sound intertidal study, the same experimental units or sites were repeatedly sampled annually, and as such, the observations are correlated through time and cannot be considered independent. Repeated measures have often been mistakenly treated as independent replicate observations in ecological studies (Hurlbert 1984). However, these dependent data violate the assumption of independence necessary for most classical univariate statistical methods. Hence, Skalski and Robson (1992, pp. 194–211) recommend using multivariate repeated measures and profile analysis to test for impact and recovery following an environmental accident. These multivariate methods properly account for the dependencies within the monitoring data.

For a multiperiod test of impact or recovery, the null hypothesis of parallelism can be written (Skalski and Robson 1992, p. 199) as:

$$H_o\colon \begin{bmatrix} \mu_{C1} - \mu_{C2} \\ \mu_{C2} - \mu_{C3} \\ \vdots \\ \mu_{C,t-1} - \mu_{Ct} \end{bmatrix} = \begin{bmatrix} \mu_{T1} - \mu_{T2} \\ \mu_{T2} - \mu_{T3} \\ \vdots \\ \mu_{T,t-1} - \mu_{Tt} \end{bmatrix}$$

against

$$H_a\colon \begin{bmatrix} \mu_{C1} - \mu_{C2} \\ \mu_{C2} - \mu_{C3} \\ \vdots \\ \mu_{C,t-1} - \mu_{Ct} \end{bmatrix} \neq \begin{bmatrix} \mu_{T1} - \mu_{T2} \\ \mu_{T2} - \mu_{T3} \\ \vdots \\ \mu_{T,t-1} - \mu_{Tt} \end{bmatrix}$$

where μ_{ij} is the mean for the ith treatment ($i = C, T$) at the jth survey period ($j = 1, \ldots, t$). Data are assumed to be transformed to the proper additive scale which will typically also stabilize the treatment variances. In the case of multivariate normal data with equal variance–covariance matrices (i.e., $\boldsymbol{\Sigma}$) for controls and treatments, a Hotellings T^2 statistic or an equivalent F test can be used to test the null hypothesis of parallelism (Morrison 1976, pp. 153–160). In the case of transitory effects, sequential tests can be performed to determine the time frames when parallelism, and hence, recovery has occurred.

Unfortunately, for multivariate ANOVA to be appropriate, more replicate sites are needed than repeated measures over time. This requirement was not satisfied with the Prince William Sound monitoring data, where the three replicate sites per category were resampled over a nine-year period. Consequently, ad hoc statistical analysis must be used to illustrate the recovery principles with the intertidal data. The P values associated with these tests of significance should be only loosely interpreted when inferring impact and recovery.

In the analysis of the Prince William Sound data, sequential testing procedures were used to test for recovery (i.e., parallelism) and to identify the periods of impact and recovery. Starting with the most recent years of data, a six-year window of time was examined. Within this time frame, a test of parallelism was performed. If the null hypothesis was not rejected, then the window was moved back one year in time and the analysis repeated. The back-step sequential procedure was continued until the null hypothesis of parallelism was rejected, indicating the period of impact. A forward sequential procedure could also be performed, beginning with the initial period of possible impact. The time window would be advanced by one-year increments until evidence of recovery is detected (i.e., null hypothesis of parallelism not rejected).

The six-year test window was used so that parametric tests would have at least four degrees of freedom for the error term. Smaller time windows are more sensitive to localized deviations from parallelism but have fewer degrees of freedom and generally lower statistical power. Longer time windows will have greater statistical power but also have inherently less temporal resolution. Any choice of window size is a compromise between these opposing forces of resolution and statistical power.

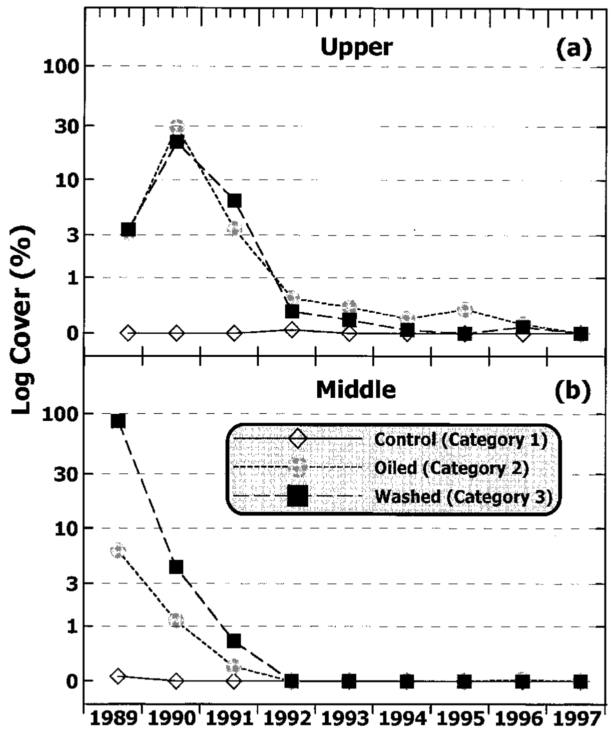

Figure 4. Time profile of average percent coverage of intertidal areas by *Exxon Valdez* oil by category for (**a**) upper transect and (**b**) middle transect locations.

Two related parametric models were used to test for parallelism within selected time windows. The first model is based on a time × treatment ANOVA that allowed all three treatments to be analyzed simultaneously. The approach was to fit a common polynomial to mean values over time and then test whether categories 1–3 showed the same temporal trends by testing for a treatment-by-time interaction. Separate intercepts were fit to each category because there was no a priori reason to assume the impact sites would recover to the mean of the control sites. Mean abundance was log-transformed prior to analysis. The second approach was to analyze the logarithm of the ratio of the means at control and treatment (i.e., impact) sites where the dependent variable was defined as

$$y_i = \log\left(\frac{\bar{x}_{Ci}}{\bar{x}_{Ti}}\right)$$

for the *i*th year ($i = 1, \ldots, 6$). Polynomial regression of the log-ratio versus time provided an additional test of parallelism, which was rejected if the regression coefficients were statistically different from zero. The polynomial regression of the log-ratio was performed on categories 2 and 3 sites separately.

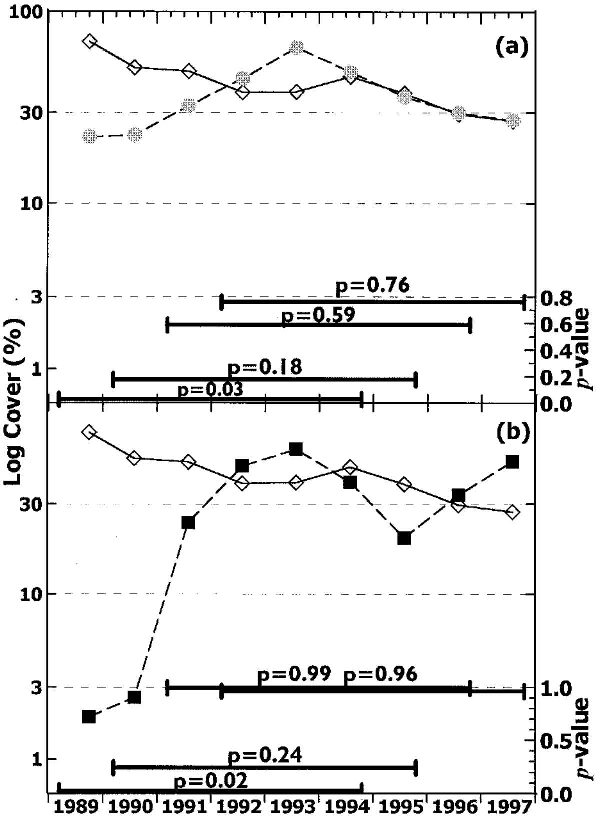

Figure 5. Comparison of temporal profiles of rockweed coverage for parallelism at the middle intertidal transects between control (i.e., category 1, denoted by ◇) and (**a**) category 2 and (**b**) category 3. Horizontal bars illustrate various 6-year time windows used in tests of parallelism and corresponding *P* values.

Results

Extensive comparison of biotic time profiles of infauna and epibiota at control, oiled, and oiled–washed sites can be found in Coats and others (1999). Selected results are presented below, illustrating the results of the ad hoc tests of parallelism and corresponding graphs of time profiles.

Oil Coverage

Percent oil coverage within the intertidal quadrates used for the epibiotic investigations is illustrated in Figure 4. Although the plots do not provide a detailed measure of oil exposure, they do provide an index of hydrocarbon exposure experienced by the epibiota. The greatest initial amount of oil coverage was along the middle intertidal transects at category 3 sites (Figure 4b), but oil persisted for approximately 2–4 years longer along the upper transects (Figure 4a).

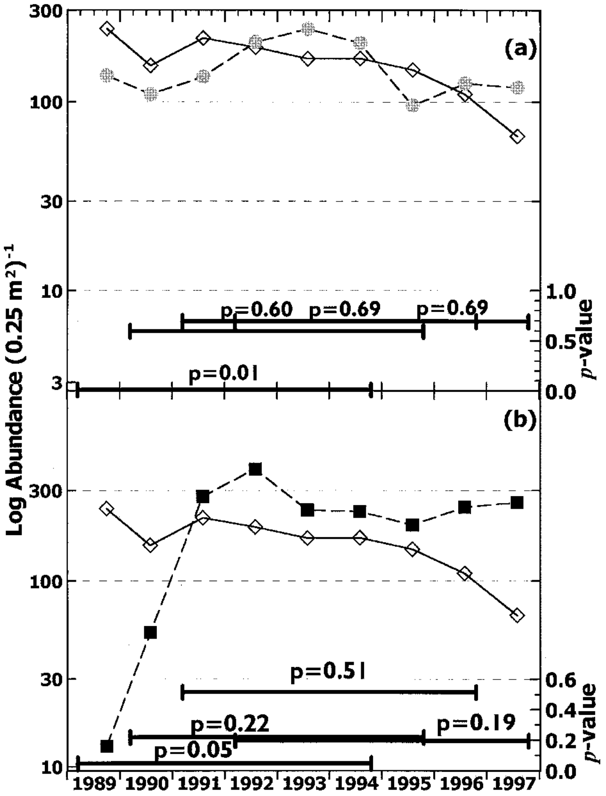

Figure 6. Comparison of temporal profiles of invertebrate abundance for parallelism at the middle intertidal transect between control (i.e., category 1, denoted by ◇) and (**a**) category 2 and (**b**) category 3. Horizontal bars Illustrate various 6-year time windows used in test of parallelism and corresponding *P* values.

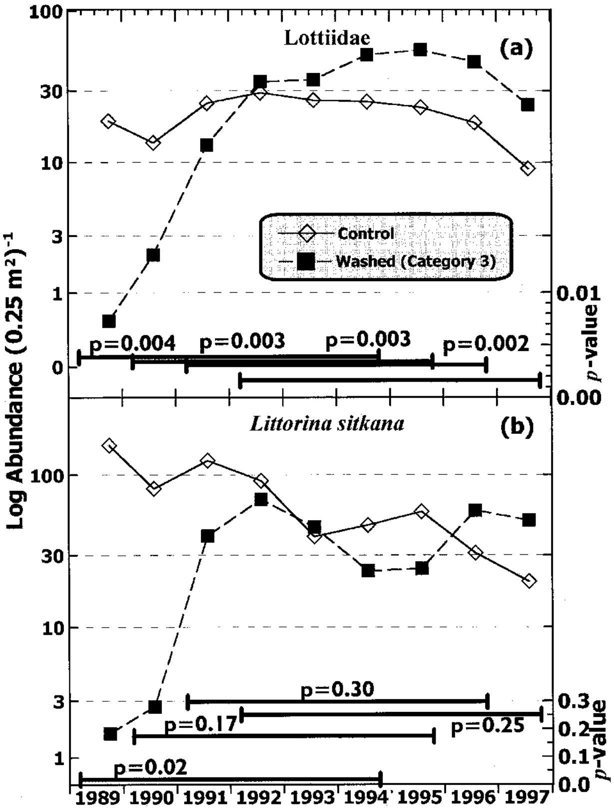

Figure 7. Comparison of temporal profiles of the abundance (**a**) limpets and (**b**) Sitka periwinkles for parallelism at the middle intertidal transects. Horizontal bars illustrate various 6-year time windows used in test of parallelism and corresponding *P* values.

Algae

Fucus, the dominant noncrustose algal species within the middle intertidal zone of impact sites, increased significantly after 1990. By 1992, *Fucus* cover had reached comparatively stable levels, similar to those of control sites (Figure 5). This stabilization period coincided with the absence of measurable oil cover in the middle intertidal zone. The time profiles for the rockweed (*Fucus gardneri*) also illustrate a case where recovery of the category 2 and 3 sites resulted in similar response amplitudes as the control sites (i.e., category 1) after 3–4 years of progressive recovery (Figure 5).

Intertidal Epibiota

The abundance of motile invertebrates exhibited a statistically significant departure from parallelism in the first six-year window at both categories 2 and 3 (Figure 6). Recovery of the middle-transect intertidal organisms was largely complete by 1991 as indicated by the high *P* values ($P > 0.10$) found in tests of parallelism. Inspection of the middle transect for intertidal invertebrates suggested a strong signature of recovery following by parallelism beginning in 1991. For the category 3 sites, recovery occurs at an abundance level above that of the controls, while for category 2 sites, recovery levels are nearly the same as at the controls (Figure 6). The null hypothesis of parallelism could not be rejected for any time window at the upper transects of categories 2 and 3 (figures not shown).

Limpets (Lottiidae) and the Sitka periwinkle (*L. sitkana*) were responsible for most of the temporal differences in epibiota abundance. Both organisms show a distinct recovery profile (Figure 7) similar to the overall invertebrate abundance. The steep early repopulation of both taxa was a major contributor to the observed nonparallelism in motile invertebrate abundance within the middle intertidal transects at category 3.

There was no statistically significant evidence of impacts on sessile epibiotoic invertebrates (i.e., barnacles and mussels). Namely, the null hypothesis that temporal trends in sessile invertebrate cover at category 2 and

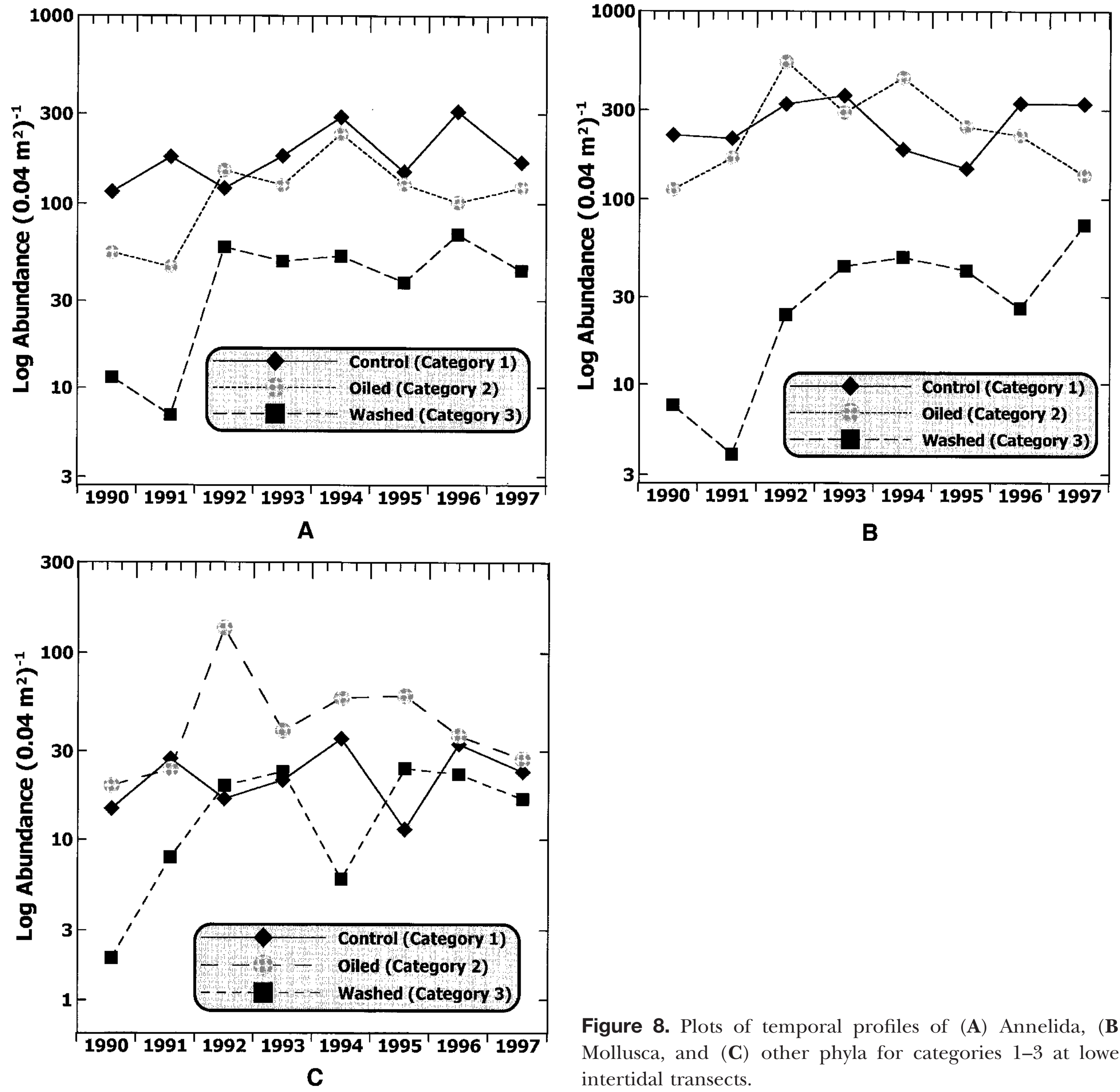

Figure 8. Plots of temporal profiles of (**A**) Annelida, (**B**) Mollusca, and (**C**) other phyla for categories 1–3 at lower intertidal transects.

3 sites and control sites were parallel could not be rejected in the initial 6-year time window. This was the case for both the upper and middle intertidal zones at both categories 2 and 3.

Intertidal Infauna

A major event within the infaunal community as a whole was the recovery of total abundance between 1991 and 1992. The recovery included diverse taxonomic groups including annelids, mollusks, and other infauna (Figure 8). Although the crustacean time profile at impacted sites exhibited a sharp population increase in 1991 (Figure 9), the recovery of the taxonomic group is not as clearly defined as other taxa with category 3 sites showing a divergent trend in recent years.

At the category 3 oiled and washed sites, the abundance of annelida (Figure 8A) and mollusca (Figure 8B) is tracking control abundance through time, suggesting recovery. However, recovery for these taxa at category 3 sites is at a lower absolute abundance level than at control sites. Tests of parallelism suggest both category 2 and 3 sites have recovered for these taxa. If tests of recovery had been based solely on comparisons of absolute abundance, differences would have been detected, leading to possibly wrong conclusions.

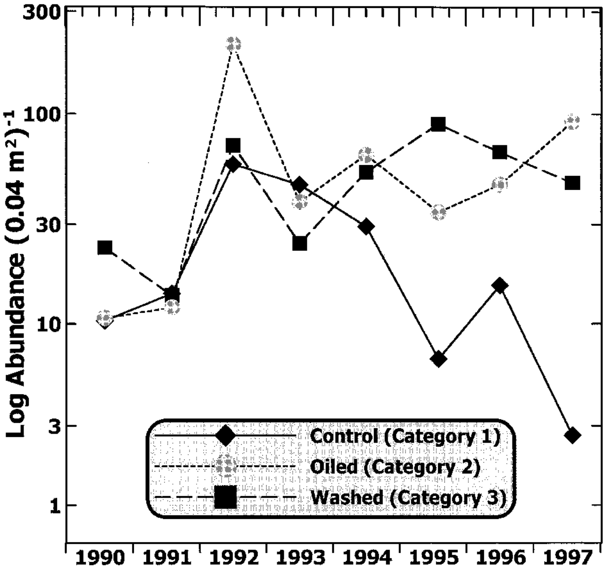

Figure 9. Plots of temporal profiles of Crustacea for categories 1–3 at lower intertidal transects.

Discussion

Recovery defined as the tracking of time profiles for control and impacted sites appears to be empirically supported by the data collected at intertidal environments following the *Exxon Valdez* oil spill. Both individual taxa and the general abundance of groups of intertidal invertebrates demonstrated initial impacts followed by returns to parallelism 2–3 years after the oil spill. This temporal pattern of apparent recovery is repeated across numerous taxa, suggesting the result is not simply an artifact of sampling. The apparent pattern of recovery is also consistent with a hypothesis of acute mortality immediately after the oil spill, followed by eventual recovery of the populations as the toxicity dissipates. The observed temporal profiles strongly support the contention that the intertidal epibiota populations have returned to levels that could have prevailed in the absence of the oil spill. The temporal profiles for the control and oiled infauna sites also suggest similar recovery. However, ambient factors such as grain size differences still persist, indicating recovery of infaunal communities may still be continuing (Hayes and Michel 1998).

Determination of whether the populations have returned to levels that would have prevailed in the absence of an impact were based on tests of parallelism of the mean temporal profiles at control and oiled sites. Examples from various taxa suggest that these recovered levels may be less than (i.e., Figure 8A, B), equal (e.g., Figure 5a, b), or greater than (e.g., Figure 7a) control levels of abundance. A simple test for the equality of abundance would have masked or obscured the occurrence of recovery for many taxa. We believe our definition of recovery has general application to many of the oil spill recovery programs following the *Exxon Valdez* oil spill (Table 1) and to many other accident assessment studies.

There may be circumstances when field programs will not have sufficient numbers of study areas for rigorous statistical testing. Reasons include topographic or regulatory constraints. In the case of the *Exxon Valdez* oil spill, there was tremendous public pressure to minimize the number of oiled but not cleaned sites (i.e., category 2). These design limitations can eventually hamper rigorous tests of impact and recovery. Skalski and Robson (1992, pp. 205–213) provide power calculations for designing accident assessment studies. The power of statistical tests is a complex function of sample size, temporal variances and covariances, and the temporal pattern of impact and recovery. At a minimum, three years of study are needed to test for impact using the approach recommended above. To confirm that recovery has occurred, the study must continue for a minimum of two years after recovery (i.e., three years of recovery). In all cases, there need to be more replicate study areas per treatment level than years of study in order to perform multivariate profile analysis. As such, oil spill studies need to be designed from the onset to accommodate the requirements for both impact assessment and recovery analysis. Otherwise, long-term monitoring programs may result in nothing more than anecdotal stories, rather than rigorous scientific investigations, of environmental perturbations caused by humankind.

Acknowledgments

This study was funded by the US Department of Commerce and National Oceanic and Atmospheric Administration (NOAA), Hazardous Materials Response Division (HazMat). The guidance and support provided by them throughout the course of this study are greatly appreciated.

We would also like to express our gratitude to other project members who have provided valuable assistance in the past year. They include Scott Kimura, Terric Klinger, David Janka, and Mandy Lindeberg for field collections and John Steinbeck for data management support. All of their contributions are greatly appreciated.

Literature Cited

Barlett, M. S. 1947. The use of transformation. *Biometrics* 3:39–52.

Coats, D. A., E. Imamura, A. K. Fukuyama, J. R. Skalski, S. Kimura, and J. Steinbeck. 1999. Monitoring of biological recovery of Prince William Sound intertidal sites impacted by the *Exxon Valdez* oil spill. NOAA Technical Memorandum NOS OR&R 1, Final report. Prepared for Hazardous Materials Response and Assessment Division, NOAA, Seattle, 73 pp. + appendices.

Eberhardt, L. L. 1978. Appraising variability in population studies. *Journal of Wildlife Management* 42:207–238.

EVOSTC (Exxon Valdez Oil Spill Trustee Council). 1999. *Exxon Valdez* oil spill restoration plan. Update on injured resources and services. January 1999. Anchorage, 24 pp.

Hayes, M. O. 1996. An exposure index for oiled shorelines. *Marine Pollution Bulletin* 3:139–147.

Hayes, M. O., and J. Michel. 1998. Evaluation of the condition of Prince William Sound shorelines following the *Exxon Valdez* oil spill and subsequent shoreline treatment, 1997 geomorphology monitoring survey. NOAA Technical Memorandum NOS ORCA 67, Volume III, Appendixes to the 1991 biological monitoring survey. Prepared for the Hazardous Materials Response and Assessment Division, NOAA, Seattle, 201 pp.

Holloway, M. 1996. Sounding out science. *Scientific American* 275(4):82–88.

Houghton, J. P., A. K. Fukuyama, D. C. Lees, H. Teas III, H. L. Cumberland, P. M. Harper, T. A. Ebert, and W. B. Driskell. 1993. Evaluation of the 1991 condition of Prince William Sound shorelines following the *Exxon Valdez* oil spill and subsequent shoreline treatment. NOAA Technical Memorandum NOS ORCA 67, Volume III, 1991 Biological monitoring survey appendices, pp. A-1-1–A-1-32.

Houghton, J. P., R. H. Gilmour, D. C. Lees, W. B. Driskell, and S. C. Lindstrom. 1997. Long-term recovery (1989–1996) of Prince William Sound littoral biota following the *Exxon Valdez* oil spill and subsequent shoreline treatment. Prepared for NOAA Hazardous Materials Response and Assessment Division, NOAA, Seattle, 130 pp. + appendices.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.

Kirk, R. E. 1982. Experimental design: Procedures for the behavioral sciences. Wadsworth, Belmont, CA, 911 pp.

Malthus, T. R. 1798. An essay on the principle of population, as it affects the future improvement of society, with remarks on the speculations of Mr. Godwin, M. Condorcet and other writers. Printed for J. Johnston in St. Paul's churchyard. Anonymous, London.

Mearns, A. J. 1996. *Exxon Valdez* shoreline treatment and operations: Implications for response, assessment, monitoring, and research. *American Fisheries Society Symposium* 18:1–16.

Michel, J., and M. O. Hayes. 1994. Evaluation of the condition of Prince William Sound shorelines following the *Exxon Valdez* oil spill and subsequent shoreline treatment, Volume II, 1997 geomorphological monitoring survey. NOAA Technical Memorandum NOS ORCA 91. Prepared for Hazardous Materials Response and Assessment Division, NOAA, Seattle, 120 pp. + appendices.

Morrison, D. F. 1976. Multivariate statistical methods. McGraw-Hill, New York.

Rice, S. D., R. B. Spies, D. A. Wolfe, and B. A. Wright (eds). 1996. Proceedings of the *Exxon Valdez* oil spill symposium, 2–5 November 1993, Anchorage, 931 pp.

Roberts, P. O., C. B. Henry, Jr., R. LeBlanc, K. Deroche, R. East, and E. B. Overton. 1997. Summary of chemistry results: Shoreline monitoring study in Prince William Sound, Alaska during the summer of 1995. Chemistry report IES/RCAT 97–42. Prepared for Hazardous Materials Response and Assessment Division, NOAA, Seattle, 50 pp. + appendices.

Skalski, J. R. 1995. Statistical considerations in the design and analysis of environmental damage assessment studies. *Journal of Environmental Management* 43:67–85.

Skalski, J. R., and D. S. Robson. 1992. Techniques for wildlife investigations: Design and analysis of capture data. Academic Press, New York, 237 pp.

Spies, R. B., S. D. Rice, D. A. Wolfe, and B. A. Wright. 1996. The effects of the *Exxon Valdez* oil spill on the Alaskan coastal environment. *American Fisheries Society Symposium* 18: 1–16.

Tukey, J. W. 1949. One degree of freedom for additivity. *Biometrics* 5:232–242.

Wells, P. G., J. N. Butler, and J. S. Hughes (eds). 1995. Exxon Valdez oil spill: Fate and effects in Alaskan waters. STP 1219. American Society for Testing and Materials, Philadelphia, 955 pp.