Making Sense of Scientists and "Sound Science": Truth and Consequences for Endangered Species in the Klamath Basin and Beyond

*Daniel J. McGarvey** *& Brett Marshall***

The National Research Council's 2002 "Interim Report on Endangered and Threatened Fishes in the Klamath River Basin" and 2004 "Endangered and Threatened Fishes in the Klamath River Basin: Causes of Decline and Strategies for Recovery" have become the collective cornerstone of what is arguably the most controversial scenario in the history of the Endangered Species Act. The Klamath Project, a federal irrigation venture that conflicts with the needs of three threatened or endangered species of fish in the upper Klamath Basin, is now the poster child for a congressional movement to reform the Endangered Species Act. This movement has acquired much of its perceived validity from misinformed interpretations of the two National Research Council reports – interpretations which the National Research Council has not challenged. To alleviate some of the confusion surrounding the reports, this Comment provides a basic introduction to the scientific method, engages in a review of the analytical process that the National Research Council used, accounts

* Ph.D. candidate in Fish Biology and Aquatic Ecology, The University of Alabama, 2006. M.S., Pennsylvania State University, 2001. B.A.,Wittenberg University, 1997. I wish to thank the Karuk Indian Tribe for bringing endangered species in the Klamath Basin to my attention, and William L. Andreen, Edgar L. Clarkson Professor of Law, The University of Alabama School of Law, who provided incentive, inspiration, and guidance for this Comment. Thanks also to Michael Cooperman, Douglas Markle, and Holly Doremus for valuable comments. Finally, I send my deepest gratitude to the staff of Ecology Law Quarterly, particularly Articles Editor Estie Manchik, for having the patience to work with an author whose knowledge of the law review literature is as callow as my own. This work was supported in part by the National Science Foundation (NSF/IGERT grant: DGE9972810).

** Ph.D. candidate in Ecology, Montana State University, 2006. M.S. in Entomology and Aquatic Ecology, Virginia Polytechnic Institute and State University, 1997. B.S., University of Wisconsin-Superior, 1990.

for several key sources of scientific uncertainty in the Klamath Basin, and remarks on the potential effects of the so-called "Endangered Species Data Quality Act of 2004." In particular, the Comment demonstrates why there is little evidence to suggest the Klamath Project is harming listed fishes, but even less evidence to suggest that it is not. And, it explains how the orthodox standards of scientific peer review can be, sensu stricto, incompatible with the proactive objectives of the Endangered Species Act.

INTRODUCTION

> "Science is increasingly criticized not because it is bad, but because it provides inadequate guidance to answer questions posed by legislatures and administrators."

- A. Dan Tarlock, *The Nonequilibrium Paradigm in Ecology and the Partial Unraveling of Environmental Law* 27 LOY. L.A. L. REV. 1121 (1994).

> "Although it has become customary, and is surely proper, to deplore the widening gulf that separates the professional scientist from his colleagues in other fields, too little attention is paid to the essential relationship between that gulf and the mechanisms intrinsic to scientific advance."

- THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS 21 (1970).

Uncertainty is not the mark of something less than "sound science." Given sufficient time and resources, science is adept at achieving a high margin of certainty. The vast majority of the scientific endeavor is, however, a gradual, systematic attempt to negotiate uncertainty – not a unilateral triumph over it. Scientists will often dedicate years of trial-and-error research to the publication of a single, conclusive result, learning as much from their mistakes as their victories. Thus, scientists understand that sound science is more of a process than an endpoint.

To prove the validity of a hypothetical proposition, the scientist must compile enough evidence to convince a jury of her peers that she is correct. In traditional, peer-reviewed circles, this will likely entail 95% certainty;[1] anything less is prone to rejection, regardless of its plausibility. Such a conservative standard ensures the overall quality of the scientific enterprise, but it is also fundamentally incongruous with precautionary initiatives, such as the Endangered Species Act (ESA),[2] in which a key assumption of scientific peer review breaks down. Scientists assume that the delays associated with further testing are prudent, and that the subjects of "conventional" science can afford to await greater certainty. But in reality, the intended beneficiaries of *proactive* legislation, such as endangered species, may not be at liberty to wait for conclusive results, because they are already suffering legitimate harm.[3] Regrettably,

---

1. While most peer-reviewed journals require 95% certainty to claim a significant result, the 95% benchmark is an institutional tautology; there is no analytically defensible rationale for requiring 95% certainty. *See* DAVID S. MOORE, STATISTICS: CONCEPTS AND CONTROVERSIES 416-18 (3d ed. 1991) (relating the history of significance testing and the 95% certainty standard). It is also a standard that the U.S. Court of Appeals for the D.C. Circuit has expressly rejected:

> Typically, a scientist will not so certify evidence unless the probability of error, by standard statistical measurement, is less than 5%. That is, scientific fact is at least 95% certain. Such certainty has never characterized the judicial or the administrative process. It may be that the "beyond a reasonable doubt" standard of criminal law demands 95% certainty. But the standard of ordinary civil litigation, a preponderance of the evidence, demands only 51% certainty.

Ethyl Corp. v. Environmental Protection Agency, 541 F.2d 1, 28 (D.C. Cir. 1976) (citations omitted).

2. 16 U.S.C. § 1531-44 (2000).

3. In the watershed *Ethyl Corp.* decision, Judge Wright further acknowledged the critical disparity between the conservative standards of scientific peer review and the mission of environmental statutes:

> While awaiting [95%] certainty may constitute the typical mode of scientific behavior, its appropriateness is questionable in environmental medicine, where regulators seek to prevent harm that often cannot be labeled "certain" until after it occurs. . . . Where a statute is precautionary in nature, the evidence difficult to come by, uncertain, or conflicting because it is on the frontiers of scientific knowledge . . . we will not demand rigorous step-by-step proof of cause and effect. Such proof may be impossible to obtain if the precautionary purpose of the statute is to be served.

scientists have rarely communicated this critically important caveat, in a transparent manner, to non-scientific audiences. This Comment attempts to rectify that oversight and facilitate more effective dialogues among scientific and non-scientific parties. To do so, the Comment introduces some alternative, but equally valid tools for dealing with uncertainty, and explains how an inexorable adherence to scientific convention recently contributed to the misappropriation of "sound science" and Endangered Species Act provisions in the upper Klamath Basin of southern Oregon and northern California.[4]

The Klamath Project (KP) is a federally subsidized reclamation venture that supplies water to almost 200,000 acres of irrigated farmland in the arid, upper Klamath Basin.[5] With total annual rainfall for this area averaging a mere 14 inches,[6] the KP is the lifeblood of approximately 1,400 privately owned farms.[7] Upper Klamath Lake (UKL) is the primary source of KP irrigation water;[8] in most years, UKL contributes between 350,000 and 450,000 acre feet of water to the KP.[9]

UKL is also home to two species of federally endangered fishes: the Lost River sucker (*Deltistes luxatus*) and shortnose sucker (*Chasmistes brevirostris*).[10] Both are large,[11] long-lived[12] species that occur only within

---

541 F.2d at 25, 28.

4. *See generally* Holly Doremus & A. Dan Tarlock, *Fish, Farms, and the Clash of Cultures in the Klamath Basin*, 30 ECOLOGY L.Q. 279, 316-335 (2003) (discussing the history of the Klamath Project and Endangered Species Act activities within the Klamath Basin); Sharon Levy, *Turbulence in the Klamath River Basin*, 53 BIOSCIENCE 315 (2003); Robert F. Service, *'Combat Biology' on the Klamath*, 300 SCIENCE 36 (2003) (providing more technical reviews of ongoing events in the Klamath Basin).

5. NATIONAL RESEARCH COUNCIL, ENDANGERED AND THREATENED FISHES IN THE KLAMATH RIVER BASIN: CAUSES OF DECLINE AND STRATEGIES FOR RECOVERY 82-83 (2004) [hereinafter FINAL REPORT], *available at* http://www.nap.edu/books/0309090970/html.

6. The annual rainfall value is a 50-year (1950-2000) average, calculated with historical data from Klamath Falls, Oregon. For reference purposes, annual average rainfall, for the same period of time, is shown parenthetically for each of the following cities: Seattle (38 inches); New York (38 inches); Atlanta (48 inches); Chicago (37 inches); Phoenix (8 inches); and Los Angeles (12 inches). All rainfall data were obtained from the National Climatic Data Center, http://www.ncdc.noaa.gov/oa/ncdc.html (last visited February 20, 2005).

7. FINAL REPORT, *supra* note 5, at 83.

8. *Id.* at 20-23.

9. *See* U.S. FISH AND WILDLIFE SERVICE, BIOLOGICAL/CONFERENCE OPINION REGARDING THE EFFECTS OF OPERATION OF THE U.S. BUREAU OF RECLAMATION'S PROPOSED 10-YEAR OPERATION PLAN FOR THE KLAMATH PROJECT AND ITS EFFECT ON THE ENDANGERED LOST RIVER SUCKER (*DELISTES LUXATUS*), ENDANGERED SHORTNOSE SUCKER (*CHASMISTES BREVIROSTRIS*), THREATENED BALD EAGLE (*HALIAEETUS LEUCOCEPHALUS*) AND PROPOSED CRITICAL HABITAT FOR THE LOST RIVER AND SHORTNOSE SUCKERS 21 (2002) [hereinafter FWS 2002 BIOP], *available at* http://klamathfallsfwo.fws.gov/hot/hot.html. The maximum capacity of UKL is 603,000 acre feet. FINAL REPORT, *supra* note 5, at 96.

10. Three ESA-listed fishes reside in the greater Klamath Basin. The Lost River sucker (*Deltistes luxatus*) and shortnose sucker (*Chasmistes brevirostris*), which were listed as federally endangered in 1988, spend their entire lives in the lakes and rivers of the upper Klamath Basin,

upper Klamath Basin lakes.[13] Adult Lost River suckers become sexually mature between 4-7 years of age,[14] while shortnose suckers begin reproduction at 4-5 years of age.[15] Spawning occurs primarily between mid-April and early May;[16] adults of both species migrate either to upstream tributaries or shallow, near-shore lake habitats (which are usually associated with freshwater springs),[17] where they will deposit between 18,000-236,000 eggs per individual in coarse, gravel substrates.[18] Upon hatching, larval Lost River and shortnose suckers move quickly into the aquatic vegetation present along the lake's shorelines, which provides both cover from predators and an abundant source of invertebrate prey.[19] Meanwhile, adult suckers have an apparent preference for deeper, open-water habitats.[20]

Ever since their formal ESA listing, the Lost River and shortnose suckers have maintained a tenuous coexistence with the KP. In 2001, concern that the KP was in violation of the ESA led the U.S. Bureau of Reclamation (USBR), which oversees the KP, to prepare a biological assessment (USBR 2001) for KP operations.[21] The U.S. Fish and Wildlife Service (FWS) replied with a biological opinion (FWS 2001 BiOp), which recommended that more water be kept in UKL than USBR 2001 had

---

above Iron Gate Dam, and are therefore overseen by the FWS. *See generally* FWS 2002 BIOP, *supra* note 9, at 22-42. The coho salmon (*Oncorhynchus kisutch*), which was listed as federally threatened in 1997, migrates from the Pacific Ocean to spawn in the tributaries of the mainstem Klamath River below Iron Gate Dam and is therefore overseen by the National Marine Fisheries Service (now designated as the "National Oceanic and Atmospheric Administration Fisheries Service"). While all three threatened or endangered fishes have been central to the ongoing scientific and political debates in the Klamath Basin, this Comment focuses solely upon the Lost River and shortnose suckers in UKL. For complete information on the status of threatened coho salmon in the Klamath River, see FINAL REPORT, *supra* note 5, at 250-310. *See also* NATIONAL MARINE FISHERIES SERVICE, BIOLOGICAL OPINION: KLAMATH PROJECT OPERATIONS (May 31, 2002), *available at* http://swr.ucsd.edu/psd/klamath/KpopBO2002finalMay31.pdf.

11. Lost River suckers can grow to 39 inches in length, while shortnose suckers generally do not exceed 20 inches. FWS 2002 BIOP, *supra* note 9, at 39.

12. Lost River and shortnose suckers are know to live upwards of 43 and 33 years, respectively. *Id.*

13. *Id.*

14. *Id.*

15. *Id.*

16. *Id.* at 40.

17. *Id.* at 39-40.

18. *Id.* at 40.

19. *Id.* at 41.

20. *Id.* at 42.

21. U.S. BUREAU OF RECLAMATION, BIOLOGICAL ASSESSMENT OF KLAMATH PROJECT'S CONTINUING OPERATIONS ON THE ENDANGERED LOST RIVER SUCKER AND SHORTNOSE SUCKER (2001), *available at* http://www.usbr.gov/mp/kbao/esa/4-1-01_30_Final_Sucker_BA.pdf.

allocated.[22] Under the ESA,[23] the USBR was obliged to embrace the FWS 2001 BiOp management alternatives. During the drought summer of 2001, however, compliance with the FWS 2001 BiOp amounted to a complete suspension of KP water deliveries from UKL.[24] This elicited a tremendous outcry from KP farmers, who suffered significant financial losses.[25] In response, the U.S. Department of Interior (DOI) commissioned a *post hoc* National Research Council (NRC) review of the FWS 2001 BiOp's scientific merit.[26]

The NRC Committee on Endangered and Threatened Fishes in the Klamath River Basin (NRC Committee),[27] which was appointed to provide a purely objective, peer-reviewed critique of the FWS 2001 BiOp,[28] quickly prepared an "Interim Report on Endangered and

---

22. *See* U.S. FISH AND WILDLIFE SERVICE, BIOLOGICAL/CONFERENCE OPINION REGARDING THE EFFECTS OF OPERATION OF THE BUREAU OF RECLAMATION'S KLAMATH PROJECT ON THE ENDANGERED LOST RIVER SUCKER (*DELTISTES LUXATUS*), ENDANGERED SHORTNOSE SUCKER (*CHASMISTES BREVIROSTRIS*), THREATENED BALD EAGLE (*HALIAEETUS LEUCOCEPHALUS*) AND PROPOSED CRITICAL HABITAT FOR THE LOST RIVER/SHORTNOSE SUCKERS 3(2) 144-154 (2001) [hereinafter FWS 2001 BIOP], *available at* http://www.usbr.gov/mp/kbao/esa/34_final_sucker_bo_4_06_01.pdf. Note that FWS 2001 BiOp and FWS 2002 BiOp are not identical documents; FWS 2002 BiOp contains revisions that supersede FWS 2001 BiOp. Reference to FWS 2001 BiOp is included herein to preserve historical and analytical accuracy, as the most controversial elements of the NRC Committee review were in reference to the FWS 2001 BiOp, not the FWS 2002 BiOp.

23. 16 U.S.C. § 1531-44 (2000).

24. FINAL REPORT, *supra* note 5, at 1.

25. Monetary damages incurred by KP farmers in 2001 were estimated at $28-35 million or higher. BILL JAEGER, ALTERNATIVE APPROACHES TO WATER MANAGEMENT IN THE KLAMATH BASIN (DRAFT) 7 (2001), *available at* http://www.klamathbasincrisis.org/pdf-files/alternatives.pdf.

26. The NRC is the principal operating agency of the National Academy of Sciences; its primary directive is to provide scientific and technological guidance to the federal government, the public, and the greater scientific and engineering communities. Complete information on the National Academy of Sciences is available at http://www4.nationalacademies.org/nas/nashome.nsf (last visited February 20, 2005).

27. The NRC Committee on Endangered and Threatened Fishes in the Klamath River Basin consisted of 12 members: William M. Lewis, Jr. (Committee Chair, University of Colorado, Boulder), Richard M. Adams (Oregon State University), Ellis B. Cowling (North Carolina State University), Eugene S. Helfman (University of Georgia), Charles D. D. Howard (Consulting Engineer, Victoria, British Columbia, Canada), Robert J. Huggett (Michigan State University), Nancy E. Langston (University of Wisconsin), Jeffrey F. Mount (University of California at Davis), Peter B. Moyle (University of California, Davis), Tammy J. Newcomb (Michigan Department of Natural Resources), Michael L. Pace (Institute of Ecosystem Studies), and J. B. Ruhl (Florida State University). FINAL REPORT, *supra* note 5, at *v*.

28. The mission of the NRC Committee was to:

> [R]eview the government's biological opinions regarding the effects of Klamath Project operations on species in the Klamath River Basin listed under the Endangered Species Act, including. . . shortnose and Lost River suckers. The committee will assess whether the biological opinions are consistent with the available scientific information. It will consider hydrologic and other environmental parameters (including water quality and habitat availability) affecting those species at critical times in their life cycles, the probable consequences to them of not realizing those

Threatened Fishes in the Klamath River Basin" (Interim Report).[29] In keeping with its federal mandate, the NRC Committee did not acknowledge the special nature of ESA provisos in any of its analyses.[30] And in its accurate, albeit terse, review the NRC Committee concluded that "there is presently no sound scientific basis for recommending an operating regime for the Klamath Project that seeks to ensure lake levels higher on average than those [in USBR 2001]."[31] Unfortunately, by neglecting to explain that there was also insufficient evidence to prove the FWS 2001 BiOp was wrong, or why, in science, *failure to demonstrate that a given theory is true does not prove it to be false* – the central thesis of this Comment – the Interim Report allowed partisan interests to besmirch the integrity of both the FWS[32] and the ESA,[33] by providing the

---

> environmental parameters, and the inter-relationship of these environmental conditions necessary to recover and sustain the listed species. . . . The tasks of the committee encompass *only the scientific and technical issues that are relevant to the endangered sucker and threatened coho species.* The committee is not charged with investigating or reporting on economic dislocation or with forecasting the economic consequences of continued implementation of flows specified in the biological opinions.

NATIONAL RESEARCH COUNCIL, SCIENTIFIC EVALUATION OF BIOLOGICAL OPINIONS ON ENDANGERED AND THREATENED FISHES IN THE KLAMATH RIVER BASIN: INTERIM REPORT 32, 10 (2002)(emphasis added) [hereinafter INTERIM REPORT], *available at* http://www.nap.edu/books/0309083249/html.

29. The Interim Report was released in January 2002, within three months of its commission. *Id.* at 3.

30. *Id.* at 10.

31. *Id.* at 3-4.

32. Upon release of the Interim Report, Interior Secretary Gale Norton criticized the FWS and National Marine Fisheries Service [hereinafter "NMFS"] BiOps:

> The National Academy of Sciences' study indicates that there were flaws with respect to critical components of the analysis in the biological opinions and assessments. Significantly, among the academy's conclusions is its finding that there was no substantial scientific foundation for requiring higher water levels in Upper Klamath Lake or higher water levels in the Klamath River.... By challenging the analysis, the study will affect our decision making process for this year and future years.

Christine Souza, *NAS Report Backs Farmers in Klamath*, AG ALERT, (2002), *available at* http://www.cfbf.com/agalert/2002/aa-020602a.htm.

33. For example, U.S. House Representative James V. Hansen stated:

> Now we learn that withholding irrigation water from those frantic farmers was completely unnecessary. This latest travesty underscores the need to reform the Endangered Species Act. We've got to ground this mammoth law in sound science and stop this appalling guesswork. As Klamath Falls reminds us, sloppy science ruins regional economies and personal livelihoods. That's unacceptable.

Oversight Hearing on "Scientific Evaluation of Biological Opinions on Endangered and Threatened Fishes in the Klamath River Basin" Before the Subcomm. on Fisheries Conservation, Wildlife and Oceans of the House Comm. On Res., 107th Cong. (2002) (statement of James V. Hansen, Chair,. Comm. on Resources). Also, House Resources Committee member Greg Walden stated that "[t]he water crisis in the Klamath Basin is a case study in the unintended ill-effects of the Endangered Species Act." *Id.* And House Resources Committee member Wally Herger asserted that "[t]he Klamath report should serve as a wake up call to all Americans that

DOI with the opportunity it needed to circumnavigate the FWS 2001 BiOp.[34]

The following year (2002), the DOI tendered the Interim Report as justification for dismissing the FWS 2001 BiOp, soliciting a revised BiOp that was more congruent with KP objectives,[35] and reinstating the KP.[36] This landmark decision catalyzed a flood of anti-ESA sentiment and provided some of the most outspoken and powerful critics of the ESA with a sympathetic congressional audience.[37] Congressman Richard W. Pombo (California) articulated many of these concerns as follows:

> The Endangered Species Act has become a program that checks species in for protection, conservation, and recovery, but never checks them out.... Moreover, numerous qualified studies assert that none of the species listed by the FWS to have been "recovered" in the United States may reasonably be claimed to have recovered as a result of the ESA. The fact is that the few recovery success stories are not even attributable to regulatory protections under the ESA, but unrelated factors such as bans on DDT and other organochlorides.... In the recent case of the Klamath Basin and the endangered sucker fish, for example, it was determined that the sucker fish needed water supplies more than the area's farmers needed it to irrigate their crops and feed their families. The result was a devastating loss of family farms, human life and economic vitality. Only after the damage was done, the National Academy of Science (NAS) determined that [the] decision by the federal government to shut off irrigation water to nearly 1,200 farmers and ranchers had "no sound scientific basis."[38]

Such resentment culminated in the "Endangered Species Data Quality Act of 2004"[39] (which was originally titled the "Sound Science for

---

sound science and community well-being are on the chopping block when it comes to making species decisions. For this reason, the ESA needs to be updated to reflect balance and common sense." *Id.*

34. *See* Secretary Norton's comment in Souza, *supra* note 32.

35. FWS 2002 BIOP, *supra* note 9.

36. INTERIM REPORT, *supra* note 28, at 3-4.

37. *See, e.g.,* the remarks of Rep. Hansen, *supra* note 33.

38. RICHARD W. POMBO, THE ESA AT 30: A MANDATE FOR MODERNIZATION (2003) (arguing that the ESA has not been a cost-effective means of conserving threatened and endangered species in the United States), *available at* http://www.crowley-offroad.com/esa_at_30.htm.

39. H.R. 1662, 108th Cong. (2003); S. 2009, 108th Cong. (2004) (submitting an identical bill to the Senate). The Act was not passed, and as of April 13, 2005, has not been reintroduced in the 109th Congress. In summary, the Data Quality Act:

> A) Amends the Endangered Species Act of 1973 to direct the Secretary of the Interior to: (1) give greater weight to scientific and commercial data that is empirical or that has been field-tested or peer-reviewed in determining that a species is an endangered or threatened species; and (2) promulgate regulations that establish criteria for data to be used as the basis of such a determination.

Endangered Species Act Planning Act," but renamed on November 19, 2004.)[40] If passed, the Endangered Species Data Quality Act of 2004 (Data Quality Act) will institute mandatory peer review, similar to that of the NRC Committee, of all ESA listings, de-listings, recovery plans, and jeopardy opinions.[41] It will also entitle the Secretary of Interior (Secretary) to personally appoint "qualified" reviewers.[42]

> B) Prohibits the Secretary from determining that a species is endangered or threatened unless the determination is supported by field data. Requires the Secretary to include data collected by landowners in the rule-making record of such a determination.
>
> C) Requires the Secretary to publish a description of additional scientific and commercial data that would assist in the preparation of a recovery plan.
>
> D) Directs the Secretary to: (1) solicit recommendations from the National Academy of Sciences and develop a list of qualified reviewers to participate in independent scientific review actions; and (2) randomly appoint from such list three individuals who shall report on the scientific information and analyses on which final action is based.
>
> E) Requires the Secretary, when consulting with a Federal agency to determine whether agency action will jeopardize an endangered or threatened species or destroy the critical habitat of such species, to: (1) consider information provided by affected States; and (2) allow any person who has sought agency authorization or funding for an action to participate in related consultations.

The Data Quality Act is only the most recent attempt to rewrite ESA statutes. At least three others have been proposed to Congress within the past decade: the "Endangered Species Conservation and Management Act of 1995," the "Endangered Species Act Common Sense Act of 2000," and the "Endangered Species Listing and Delisting Process Reform Act of 2003." Like the Data Quality Act, each of these bills sought to impose explicit, and ultimately more conservative, criteria for receiving the benefit of ESA protection. As of April 13, 2005, however, none has been enacted. *See generally* Michael J. Brennan at al., *Square Pegs and Round Holes: Application of the "Best Scientific Data Available" Standard in the Endangered Species Act*, 16 TUL. ENVTL. L.J. 387, 433-441 (2003) (discussing the recent history of congressional movements to reform the ESA).

40. Pending its November 19, 2004 placement on Union Calendar No. 479 in the U.S. House of Representatives (108th Congress, 2d Session), the "Sound Science for Endangered Species Act Planning Act" was renamed the "Endangered Species Data Quality Act." The language of the Data Quality Act is, however, virtually identical to that of its predecessor.

41. H.R. 1662, 108th Cong. § 3(j)(1)(A) (2004).

42. The Data Quality Act defines a "qualified individual" as:

> [A]n individual with expertise in the biological sciences (i) who through publication of peer-reviewed scientific literature or other means, has demonstrated scientific expertise on the species or a similar species or other scientific expertise relevant to the decision of the Secretary...; (ii) who does not have, or represent any person with, a conflict of interest with respect to the determination that is the subject of review; (iii) who is not a participant in any petition or proposed or final determination before the Secretary; and (iv) who has no direct financial interest, and is not employed by any person with a direct financial interest, in opposing the action under consideration.

Id. at § 3 (j)(1)(B). In theory, the Data Quality Act would shield the scientific peer-review process from partisan agendas, because the appointed reviewers must be randomly selected from a list of qualified applicants. *Id.* at § 3 (j)(3). However, the right to develop and maintain lists of qualified reviewers, based upon solicited recommendations from the National Academy of

In March 2004, the NRC Committee released a more comprehensive report: "Endangered and Threatened Fishes in the Klamath River Basin: Causes of Decline and Strategies for Recovery" (Final Report).[43] The Final Report details an impressive assessment of the Klamath Basin's environmental ills, but it does not contest the DOI's suspect use of the Interim Report,[44] nor does it attempt to clarify why the conservative axioms of conventional scientific peer review, as employed by the NRC Committee,[45] are inherently antithetical to the proactive spirit of the ESA.[46] Ergo, this Comment addresses both shortcomings.

Part I of the Comment reviews the most essential, contested content of the FWS 2001 BiOp, recreates and critiques the NRC Committee's analysis, and offers a more equitable assessment, including quantitative measurements of the evidence for and against the FWS 2001 BiOp. Specifically, it demonstrates that, while scientists are only 32% certain the KP is having a significant, negative effect on endangered fishes in UKL, *they are only 2% certain that it is not*.[47] Part II outlines the basic processes of statistical inference testing, and shows how disparate, but equally sound, interpretations of the same data were reached in the upper

---

Sciences, could potentially confer significant discretionary power to the Secretary. *Id.* at § 3 (j)(2).

43. The Final Report, which was initially scheduled for completion in March 2003, was officially released in draft form October 2003 in order to solicit comments from interested parties. The National Academies Press did not release the bound, Final Report to the general public until March 2004. *See* FINAL REPORT, *supra* note 5.

44. In response to an onslaught of criticisms, both from scientific peers and ESA supporters, the Chair of the NRC Committee stated: "The Committee has no control over the uses to which its report might be put." Levy, *supra* note 4, at 319. *See, e.g.,* Michael S. Cooperman & Douglas F. Markle, *The Endangered Species Act and the National Research Council's Interim Judgment in Klamath Basin*, FISHERIES, Mar. 2003, at 10 (providing an example of a technical critique of the Interim Report). *See also* KLAMATH BASIN COALITION, NATIONAL RESEARCH COUNCIL INTERIM REPORT: SCIENTIFIC EVALUATION OF BIOLOGICAL OPINIONS ON ENDANGERED AND THREATENED FISHES IN THE KLAMATH RIVER BASIN (2003) (providing an example of an Interim Report critique that was written for the general public), *available at* http://www.klamathbasin.info/NRCfacts.pdf.

45. *See* FINAL REPORT, *supra* note 5, at 33-37, 311-316 (defending the NRC Committee's conservative assessment of the FWS 2001 BiOp).

46. 16 U.S.C. § 1536(a)(2) (2000).

> Each Federal agency shall . . . insure that any action authorized, funded, or carried out by such agency . . . is not likely to jeopardize the continued existence of any endangered species or threatened species or result in the destruction or adverse modification of habitat of such species . . . .

Id.

47. The 32% and 2% certainty statements are based upon *Mann-Whitney* and *power analysis* tests of 1990-1999 chlorophyll *a* concentration data (higher chlorophyll *a* concentrations are detrimental to water quality in UKL and to the endangered suckers), taken from the Interim Report. INTERIM REPORT, *supra* note 28, at 18. The data were divided into two test groups, in accordance with the FWS 2001 BiOp recommendation for 4141 feet of water in August; the two groups were then compared directly, allowing us to quantify the evidence for and against the FWS 2001 BiOp. *See infra* text accompanying notes 114-115.

Klamath Basin. Part III explains why the Data Quality Act will undermine, rather than strengthen, the ESA. Finally, this Article concludes that more effective management of endangered species will require scientific and non-scientific authorities to acknowledge the congenital limitations of science.

I. DIFFERENCES OF SCIENTIFIC OPINION IN THE UPPER KLAMATH BASIN

In *Fish, Farms, and the Clash of Cultures in the Klamath Basin*, Doremus and Tarlock state that "[s]cience seeks truth through a continual process of experimentation and re-evaluation. Scientists are most comfortable giving answers as ranges of probability rather than absolute causal relationships."[48] While the scientist's penchant for cryptic probability statements may initially seem peculiar, the cautious meter of scientific intercourse is, in fact, a carefully premeditated tool. At the conclusion of a study, a scientist will be expected to answer one of two questions: "are you certain" or "how certain are you"? This is a nontrivial distinction. The first question, which is the primary mechanism of peer review,[49] requires only a simple, binary response – yes or no. It therefore deprives audiences external to the research process of the opportunity to assess the weight of evidence on their own. The second approach, however, is empowering; by explicitly presenting a percent certainty, or probability value, scientists can provide their audiences with the information they need to form their own opinions. The statements "I am not certain" and "I am less than 95% certain, but more than 50% certain" convey two very different messages.

In each of the Interim and Final Reports, the NRC Committee adopted the first approach, couching its conclusions in relatively facile language, and omitting much of the technical detail that constituted the foundation of its review.[50] To impress the value of that missing detail, and

48. Doremus & Tarlock, *supra* note 4, at 325.

49. Though 95% certainty is not an explicit, literal criterion for peer-reviewed publication, the scientific community overwhelmingly recognizes it as the implicit expectation. Hence, many authors will abstain from submitting their research to peer-reviewed journals until they have collected sufficient data to claim a 95% certain result. *See* MOORE, *supra* note 1, at 416-18.

50. For example, the NRC Committee challenged and rejected the FWS 2001 BiOp claim that a fundamental relationship exists between lake elevation in UKL and a suite of water quality parameters, such as algal density:

> No relationship between lake levels and population densities of algae (as shown by chlorophyll) is evident . . . in the 9-year water-quality monitoring record that has been fully analyzed. . . . Thus, the committee concludes that there is presently no sound scientific basis for recommending an operating regime for the Klamath Project that seeks to ensure lake levels higher on average than those occurring between [the years of study].

INTERIM REPORT, *supra* note 28, at 17, 3-4. Moreover, neither the INTERIM REPORT nor the FINAL REPORT mention the outcome of a single statistical test or cite a formal percent certainty how its inclusion would have altered the principal message of the Interim and Final Reports, this Comment will revisit the NRC Committee's review of the FWS 2001 BiOp, beginning with an introduction to ongoing water quality concerns in UKL, which were central to the FWS 2001 BiOp recommendations for the KP.[51] It will then subject the available data to greater scrutiny than the NRC Committee, and provide a more comprehensive, balanced result.

A. Water Quality in UKL and the FWS 2001 BiOp

Water quality impairment represents a particularly dire threat to endangered fishes in UKL.[52] The annual proliferation of phytoplankton, or algae, within the lake is of special concern; in most years, a massive summer algal bloom, followed by the ensuing decay of dead algal tissue, substantially depletes the concentration of dissolved oxygen in UKL.[53] Given the shallow depth of UKL[54] and the warm summer air temperatures, algal blooms are prone to occur naturally,[55] and native fishes, including the endangered suckers, are largely adapted to low dissolved oxygen concentrations.[56] However, abnormally large algal blooms, the specific causes of which are uncertain, have been a persistent problem since the 1960's.[57] Under these degraded conditions, dissolved

---

value. Rather, the substantial volume of data that was considered by the NRC Committee is presented exclusively in graphical and tabular formats. *See generally id*; FINAL REPORT, *supra* note 5. Graphs and tables are useful means of summarizing data, but would generally be considered inadequate to satisfy the analytical expectations of peer-reviewed publication.

51. *See* FWS 2001 BIOP, *supra* note 22, at § 3(2) 59-97.

52. The NRC Committee recognized three water quality parameters of particular concern in UKL: dissolved oxygen, ammonia, and pH. All three are known to have direct effects on the health and survival of endangered suckers. Hence, the Committee was prepared to embrace the FWS 2001 BiOp recommendation for higher lake elevations in UKL, in the event that any of these three parameters were shown to be directly related to lake elevation. *See* FINAL REPORT, *supra* note 5, at 102.

53. Summer algal blooms within UKL are composed predominately of a single species of blue-green algae: *Aphanizomenon flos-aquae*. *Id.* at 102, 117-22 (describing the general processes within UKL that promote algal growth, death, and decay, thereby depleting dissolved oxygen). *See generally* ROBERT G. WETZEL, LIMNOLOGY: LAKE AND RIVER ECOSYSTEMS 154-68, 358-90 (3d ed. 2001) (providing more complete information on dissolved oxygen dynamics and algal life histories in lake environments).

54. The average depth of UKL is 9 feet. Maximum depth is 31 feet. FINAL REPORT, *supra* note 5, at 96-97.

55. Algal growth, which is a photosynthetic process, is limited by the availability of three essentials: light, warm water (generally speaking, algal growth accelerates with increasing water temperature), and nutrients (particularly phosphorous and nitrogen). *See* WETZEL, *supra* note 53, at 376-90.

56. FINAL REPORT, *supra* note 5, at 200.

57. Aquatic biologists typically use chlorophyll *a*, an easily measured photosynthetic pigment found in algae, as a surrogate measure of algal growth. Chlorophyll *a* concentrations in excess of 100 micrograms per liter of water are frequently observed in UKL. Such extreme concentrations are due to enhanced phosphorous availability. The specific mechanisms driving

oxygen may drop to concentrations less than the suckers' minimum tolerance level of 1-2 milligrams per liter of water, creating a lethal environment for both species of endangered suckers.[58] Low dissolved oxygen was the cause of consecutive mass mortality events, or "fish kills," in 1995, 1996, and 1997, which claimed substantial numbers of Lost River and shortnose suckers.[59] In fact, the 1996 fish kill is believed to have eliminated as much as 50% of the two species' adult populations.[60]

Because water quality degradation is known to have such severe consequences for the two endangered fishes in UKL, the FWS 2001 BiOp sought to mitigate summer algal blooms by maintaining higher than average lake elevations (i.e., larger total water volumes), specified as discrete monthly minimum elevations.[61] Higher lake elevations had the potential to minimize algal density within UKL in any of three ways – by diluting phosphorous and light availability, minimizing wind driven mixing of the vertical water column, and minimizing agitation and suspension of lake-bottom sediments – each of which will be discussed in turn.[62] First, higher lake elevations would tend to dilute phosphorous concentrations, which are proportional to algal concentrations; because phosphorous is a vital, limiting nutrient for algal growth, algal concentration is often a direct product of phosphorous availability.[63] Likewise, light availability, which is a prerequisite for algal growth, decreases at depth, so that deeper lakes support lower total algal

---

phosphorous loading in UKL have not yet been determined, however. *Id.* at 103-13. *See generally* WETZEL, *supra* note 53, at 269-86 (describing human-induced effects on phosphorous cycling in lakes and consequent effects on algal growth).

58. FINAL REPORT, *supra* note 5, at 117.

59. *See* PERKINS ET AL., THE ROLE OF POOR WATER QUALITY AND FISH KILLS IN THE DECLINE OF ENDANGERED LOST RIVER AND SHORTNOSE SUCKERS IN THE UPPER KLAMATH LAKE 11-13, 27-28 (2000), *available at* http://www.usbr.gov/mp/kbao/esa/Fish_Kill.pdf.

60. FINAL REPORT, *supra* note 5, at 238.

61. The FWS 2001 BiOp UKL elevation requirements were 4141.0 feet on January 1, 4141.5 feet on February 15, 4142.0 feet on March 15, 4142.5 feet on April 15 and June 1, 4141.5 feet on July 15, 4141.0 feet on August 15, 4140.5 feet on September 15, and 4140.0 feet on October 15. *See* FWS 2001 BIOP, *supra* note 22, at § 3(2) 143.

62. *See* FINAL REPORT, *supra* note 5, at 112; *see also* EUGENE B. WELCH & TOM BURKE, INTERIM SUMMARY REPORT: RELATIONSHIP BETWEEN LAKE ELEVATION AND WATER QUALITY IN UPPER KLAMATH LAKE, OREGON §3 15-22 (2001) (relating the modeling procedure originally used to indicate that a relationship existed between lake elevation and algal growth); FWS 2001 BIOP, *supra* note 22, at § 3(2) 72-74 (describing the general rationale for expecting higher lake elevations to confer enhanced water quality in UKL); IMST REVIEW OF THE USFWS AND NMFS 2001 BIOLOGICAL OPINIONS ON MANAGEMENT OF THE KLAMATH RECLAMATION PROJECT AND RELATED OPERATIONS 33-45 (2003) (providing an independent review and summary of multiple Klamath reports on UKL water quality and endangered suckers, including the FWS 2001 BiOp and the Interim Report), *available at* http://www.fsl.orst.edu/imst/reports/2003-01.pdf.

63. *See* FINAL REPORT, *supra* note 5, at 103-08 (describing the dynamics of phosphorous availability within UKL); *see also* WETZEL, *supra* note 53, at 275-78, 281 (describing the relationship between phosphorous availability and algal growth).

concentrations than shallower lakes.[64] Thus, as lake elevation increases, algal concentration should decrease. Second, wind-driven mixing of the vertical water profile is less pronounced in deep lakes, due to the greater frictional resistance among successive (i.e., vertical) water layers.[65] Algal decay and dissolved oxygen depletion occur primarily in the deeper waters of UKL, due to the impaired light availability and consequent algal death at depth. As a result, dissolved oxygen depletion is less likely to affect the endangered suckers, which are capable of shifting to mid-depth habitats,[66] when a deep water column is opposing the upwelling of deep, oxygen-depleted waters.[67] Finally, the inhibited potential for wind-driven mixing of the vertical water column in deeper lakes should also minimize agitation and suspension of the phosphorous-rich sediments at the bottom of UKL, which are of natural, geologic origin.[68] This is particularly important because disturbance of the water column, with subsequent suspension of bottom sediment, is the single most efficient means of increasing phosphorous concentration in UKL.[69]

B. *The NRC Committee's Failed Search for a General Relationship between Lake Elevation and Algal Growth*

In its review of the FWS 2001 BiOp, which was predicated on the above mechanisms, the NRC Committee used empirical data to search for a relationship between UKL elevation and algal density.[70] It did so by plotting a series of nine chlorophyll *a* concentration measurements, which provide a surrogate index of algal growth,[71] against their corresponding

64. *See* FINAL REPORT, *supra* note 5, at 112. *See also* WETZEL, *supra* note 53, at 56-63, 67-68 (describing the general relationship between light availability and depth and the response of algae to light availability).

65. *See generally* WETZEL, *supra* note 53, at 102-13 (describing the effects of wind on the internal circulation dynamics of lakes).

66. Adult Lost River and shortnose suckers are thought to spend the majority of their time at depths between 3-15 feet. *See* FWS 2001 BIOP, *supra* note 22, at § 3(2) 108.

67. *See generally* FINAL REPORT, *supra* note 5, at 117-20 (describing the vertical stratification of dissolved oxygen profiles in UKL).

68. The concentration of phosphorous buried in UKL sediments is approximately 25 times the concentration that is normally suspended in the lake's waters. This enormous phosphorous store is primarily a product of the local, phosphorous-rich geology, rather than human activity, such as fertilizer runoff. *Id.* at 103-105.

69. *Id.* at 106.

70. The NRC Committee used nine years of simultaneous measurements (1990-1998) to model the relationship between chlorophyll *a* and lake elevation. These data were August averages; August is a critically dry month in the Klamath Basin, when the endangered suckers are most vulnerable to dissolved oxygen depletion. *Id.* at 110-13. The data were collected by a private consulting firm (R2 Resource Consultants, Inc.) and, as of the 2001 Interim Report, represented the most complete, up-to-date UKL water quality records available. *See* WELCH & BURKE, *supra* note 62, at §3 1-5.

71. Chlorophyll *a* is the primary photosynthetic pigment of all oxygen-producing, photosynthetic organisms, such as plants and algae. It permits the conversion of solar energy to

lake elevations, and observing whether chlorophyll *a* concentration did, in fact, tend to decrease with increasing elevation.[72] This technique, referred to as regression analysis, is a statistical modeling procedure, used to track and predict how changes in one variable tend to affect a second, correlated variable.[73]

This Comment assesses the NRC Committee's analysis, and determines whether its conclusion, that "[N]o relationship between lake levels and population densities of algae is evident,"[74] is both warranted and sufficient, by reexamining the chlorophyll *a* concentration and lake elevation data. Because the NRC Committee elected to use a qualitative, graphical approach (in lieu of the more formal, quantitative procedure discussed below),[75] the first step is to identify what a graph of chlorophyll *a* concentration and lake elevation would or would not look like if, in accordance with the FWS 2001 BiOp, higher lake elevations cause algal densities to decline in UKL, thereby benefiting endangered suckers.

---

chemical, biologically available energy (i.e., glucose). Because chlorophyll *a* is relatively easy to measure empirically, it is commonly used as a surrogate for algal growth; direct, reliable measurements of algal growth are, by comparison, significantly more difficult to obtain. *See* WETZEL, *supra* note 53, at 332, 597.

72. INTERIM REPORT, *supra* note 28, at 18.

73. *See generally* MOORE, *supra* note 1, at 257-64, 281-89 (providing a general overview of statistical correlation, prediction, and regression analysis); JOHN NETER ET AL., APPLIED LINEAR REGRESSION MODELS (3d ed. 1996) (providing a comprehensive review of linear regression methods).

74. The complete opinion of the NRC Committee, regarding the FWS 2001 BiOp assertion that higher lake elevations would benefit endangered UKL suckers, was:

> Control of phosphorus in Upper Klamath Lake offers the potential of suppressing population densities of algae, thus improving water quality in the lake. No relationship between lake levels and population densities of algae (as shown by chlorophyll) is evident, however, in the 9-year water-quality monitoring record that has been fully analyzed. Thus, the idea of relieving [excessive algal growth] through phosphorus dilution caused by higher lake levels is not consistent with the irregular relationship between chlorophyll and lake level. Also, lake level fails to show any quantifiable association with extremes of dissolved oxygen or pH. For example, the most extreme pH conditions recorded for the lake over the past 10 years occurred in 1995 and 1996, which were years of intermediate water level, and not in 1992 and 1994, when water levels were lowest. (These two years had the lowest recorded water levels since 1950.) Furthermore, a substantial mass mortality occurred in 1971, the year of highest recorded water levels since 1950, and within the last ten years, mortality of adults was highest in 1995, 1996, and 1997, none of which were years of low water level. The absence of notable adult mortality in any year of low water during the 1990s might in fact suggest an association the reverse of the one postulated in the biological opinion, although the evidence is statistically inconclusive. The USFWS itself has found no association of mass mortality with lake levels. Intensified [algal growth] now affects the characteristics of the lake every year, and thus may constitute a threat to the suckers regardless of interannual variation in water level.

INTERIM REPORT, *supra* note 28, at 17 (internal citations omitted).

75. *Id.* at 18 fig.4; *see also* FINAL REPORT, *supra* note 5, at 113 fig.3-6.

These hypothetical graphs will then be contrasted with the actual, empirical data that the NRC Committee analyzed.

A graph of the relationship between chlorophyll *a* concentration and lake elevation might reasonably be expected to approximate one of the six patterns depicted in Figure 1.[76] Graph **a.)** shows a strong, inverse relationship between chlorophyll *a* concentration and lake elevation; as lake elevation increases, chlorophyll *a* concentration decreases. This relationship would make a compelling argument for the FWS 2001 BiOp recommendation that more water be left in UKL.[77] Graph **b.)** shows the exact opposite trend; as lake elevation increases, so does chlorophyll *a* concentration. This result would suggest that higher lake elevations are associated with higher algal densities and, contrary to the FWS 2001 BiOp, may actually be detrimental to endangered suckers in UKL. Graphs **c.)** and **d.)** introduce the concept of "threshold" elevations. Graph **c.)** shows that there is an inverse relationship up until some particular lake elevation, at which point elevation ceases to have such a pronounced effect upon chlorophyll *a* concentration. (This point, known as the "knee-of-the-curve," is indicated by the vertical, dotted line). Graph **d.)** illustrates a more acute threshold elevation; while chlorophyll *a* concentration and lake elevation are not correlated with each other in a one-to-one relationship, some other factor, which is associated with a specific lake elevation (indicated by the vertical, dotted line), has an obvious effect on chlorophyll *a* concentration. Both graphs **c.)** and **d.)** corroborate the FWS 2001 BiOp. Graph **e.)** suggests that no relationship exists between chlorophyll *a* concentration and lake elevation; as elevation increases, chlorophyll *a* concentration remains virtually constant, thereby contradicting the FWS 2001 BiOp. Finally, graph **f.)** is inconclusive; a correlation may exist between chlorophyll *a* concentration and lake elevation, but the data are too scattered to be certain. Thus, while graph **f.)** fails to provide support for the FWS 2001 BiOp, neither does it conclusively disprove the FWS 2001 BiOp.

---

76. While the six patterns depicted in Figure 1 are by no means an exhaustive catalog of the possible chlorophyll *a* and lake elevation relationships, they provide an adequate representation of the linear relationships that could realistically be assessed with only nine data points. With the benefit of a larger dataset it would be possible to investigate more complex, non-linear relationships. *See generally* NETER ET AL., *supra* note 73, at 95-143, 531-630 (providing a detailed explanation on higher-order, non-linear regression models). The patterns in Figure 1 are, however, adequate for the illustrative purposes of this Comment.

77. *See* FWS 2001 BIOP, *supra* note 22, at § 3(2) 72-74.

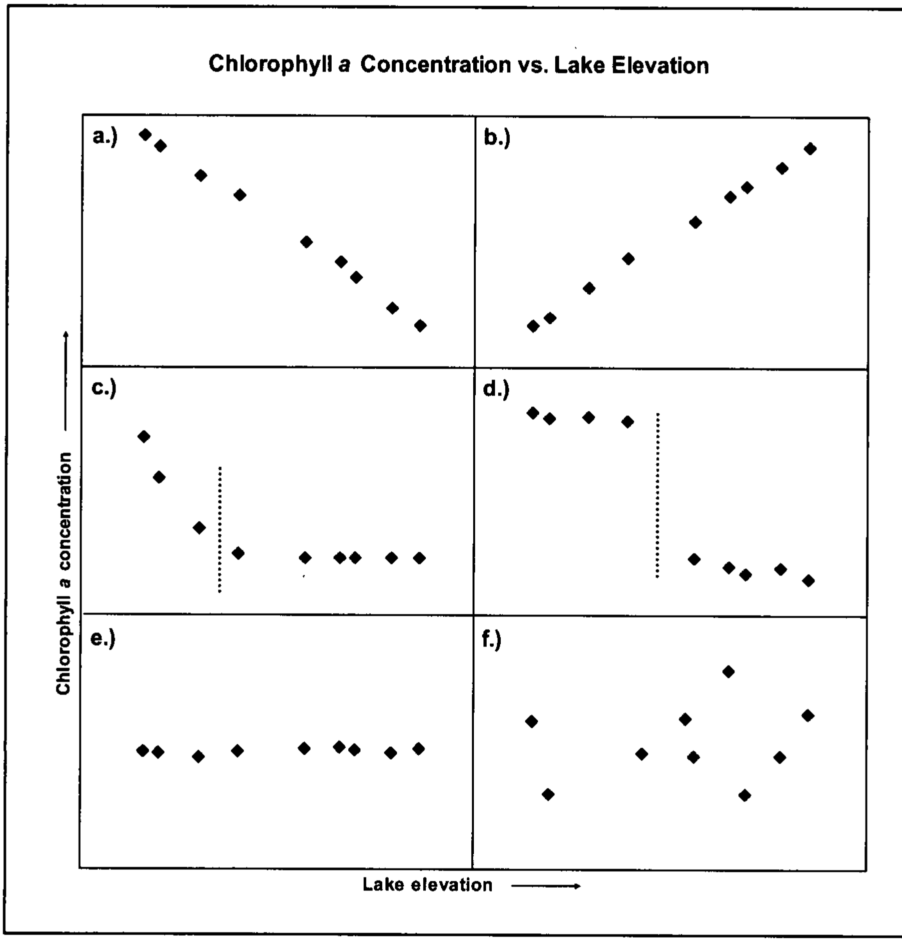

Figure 1. Six graphs of the potential relationship between chlorophyll *a* concentration (vertical axis) and lake elevation (horizontal axis) in Upper Klamath Lake. Lake elevation increases at a constant rate from left to right, while chlorophyll *a* concentration increases at a constant rate from bottom to top. Vertical, dotted lines in **c.)** and **d.)** indicate threshold elevations, beyond which the relationship between chlorophyll *a* concentration and lake elevation changes. Graph **f.)** is a plot of the actual chlorophyll *a* concentration and lake elevation data.

Had the chlorophyll *a* concentration and lake elevation data resembled any of graphs **a.)** through **e.)**, the NRC Committee's task would have been simple, and formal statistical procedures would have contributed little to their analysis. Unfortunately, graph **f.)** is a plot of the actual data.[78] Upon examining this graph, the NRC Committee concluded that chlorophyll *a* concentration and lake elevation were not correlated.[79] It was, however, remiss in not supporting that assertion with formal, quantitative statistics, despite the fact that such statistics were readily available. This was an imprudent oversight, because such statistics could have altered the legacy of the Interim and Final Reports. This Comment has therefore taken the liberty of performing such a quantitative, statistical analysis.

Formal, quantitative regression analysis, which the NRC Committee did not employ in the Interim Report, or in the Final Report (the NRC Committee applied only a visual, qualitative technique), accomplishes two objectives. First, it uses a mathematical formula called the "least-squares" algorithm to draw the "best fit" regression line between two variables.[80] Figure 2 depicts the best fit regression line for the nine years of chlorophyll *a* concentration and lake elevation data that the NRC Committee examined. This line simultaneously minimizes the sum of the squared vertical distances (short, dotted lines) between the regression line (solid, diagonal line) and each of the individual data points.[81] In this context, formal regression analysis is important because it provides a standardized, unbiased method for modeling the relationship between chlorophyll *a* and lake elevation.[82] It is worth noting, however, that in many instances, the best fit regression line can be visually interpolated with a moderately high degree of precision.[83] Thus, the trajectory of the best fit regression line, relative to a perceived relationship between chlorophyll *a* concentration and lake elevation, does not provide exceptional information; by simply looking at the graph, one may determine that, contrary to the FWS 2001 BiOp assertion that higher lake elevations would reduce chlorophyll *a* concentrations, higher lake elevations actually seem to be weakly correlated with *higher* algal densities. But unlike the visual, qualitative procedure used by the NRC Committee, formal regression analysis provides a second, equally crucial piece of information.

---

78. INTERIM REPORT, *supra* note 28, at 18 fig.4.

79. *See supra* note 74.

80. MOORE, *supra* note 1, at 283-84 (describing the "best fit" regression line).

81. Because some of the data points will always lie above the regression line, while others will always lie below it, squared distances are used to remove the effects of positive and negative values; when squared, all distances become positive values. *See id.* at 284. Using these data, the equation for the best fit regression line between chlorophyll *a* concentration and UKL elevation is: chlorophyll $a = (7.01 \times \text{lake elevation}) - 28886$.

82. MOORE, *supra* note 1, at 283-84

83. *Id.* at 281-83.

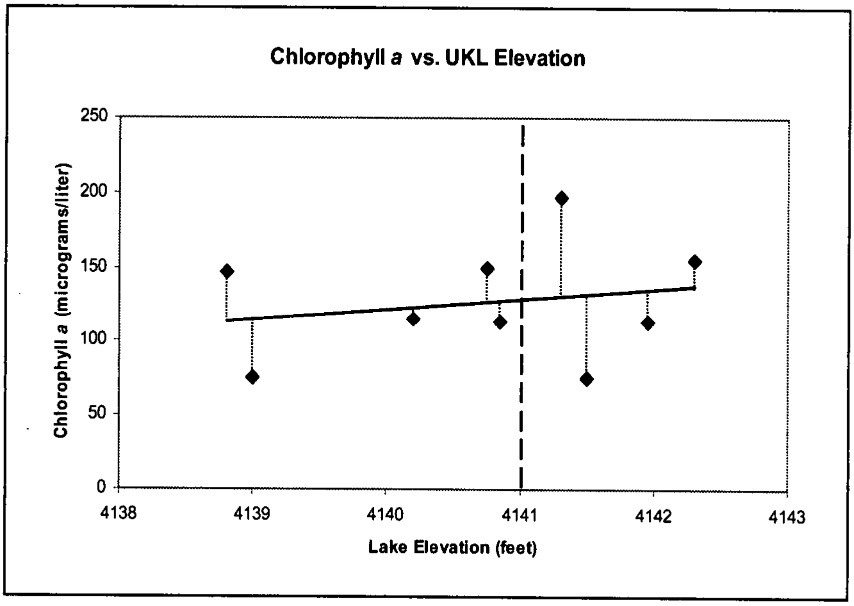

Figure 2.[84] Graph of the average August chlorophyll *a* concentrations and lake elevations in Upper Klamath Lake (1990-1998). Chlorophyll *a* data are shown as diamonds. The best fit, or least-squares regression line is the solid, diagonal line. The small, dotted vertical lines are the distances between the regression line and the actual, observed chlorophyll *a* values. The large, dashed vertical line separates the data into years when average August lake elevation was below the minimum FWS 2001 BiOp recommendation (4141 feet), and years when elevation was at or above the prescribed minimum.[85]

Once the best fit regression line has been identified, formal regression analysis can be used to assess the overall quality of that line. The least-squares algorithm ensures that the best fit regression line is the "best one possible," but it does not indicate whether that line is "good enough." Formal regression analysis solves this problem by using two statistics to test the quality of the best fit regression line: 1.) the coefficient of determination, which is a systematic measure of the *straight-line strength of association* between two variables, such as chlorophyll *a* concentration and lake elevation (r^2);[86] and 2.) the

84. Adapted from INTERIM REPORT, *supra* note 28, at 18 fig.4.

85. For the specific figures, see *supra* note 61.

86. Because r^2 values always range between 0 (i.e., zero correlation) and 1 (i.e., perfect correlation), they are most easily expressed as percentage values. The value r^2 is literally "the proportion of the variance of one variable that can be explained by straight-line dependence on

probability that this association, or correlation, is due to mere chance (*P*-value), rather than a fundamental, causal relationship between those variables.[87] Together, these two statistics can provide a rigorous, quantitative measure of the quality of the regression relationship,[88] allowing scientists to determine when the best fit regression line is a useful model of the relationship between two variables, and when it is not.

As it turns out, formal regression analysis provides very little evidence for a significant correlation, positive or negative, between chlorophyll *a* concentration and water elevation in UKL; only 5% (r^2 = 0.05) of the total variation among the nine chlorophyll *a* concentration observations is accounted for by lake elevation.[89] That is, changes in lake elevation do not seem to be associated with concurrent changes in chlorophyll *a* concentration – at least not in a straight-line, linear manner.[90] More importantly, the probability of the best fit regression line (which we already know is not exceptionally trustworthy, due to the low r^2 value) being a result of random variation, and, consequently, not a meaningful indicator of the "true" relationship between chlorophyll *a* and lake elevation, is 58% (*P*-value = 0.58).[91] Therefore, contrary to the NRC

---

the other variable." MOORE, *supra* note 1, at 262. Alternatively, one may choose to think of r^2 as the degree to which two variables change at the same *rate*.

87. There is always a possibility, however slim, that two completely unrelated variables will appear to be highly correlated. For example, suppose that five people flip a coin 10 times apiece, recording the number of "heads" and "tails" that each person obtains. Now assume that the ages of the five test subjects are 10, 20, 30, 40, and 50 years, and that the number of "head" results for each subject turns out to be 1, 4, 6, 8, and 9, respectively. A plot of the least-squares regression line for these data will be remarkably straight, and the association between age and number of heads is very strong: age accounts for 97% of the variability within the coin flipping results. These results are *highly* unlikely, and it would be ridiculous to conclude that age and coin-flipping ability are somehow correlated (as opposed to, say, writing ability). Formal regression analysis includes an estimate of the probability that an observed association, regardless of its strength, may, in fact, be spurious. This estimate is the *P*-value. In this particular instance, the probability of the coin flipping results being due entirely to chance is about 1 in 1000, or 0.1% (p = 0.001). Abysmal odds, but possible. The *P*-value is the primary means by which scientists determine whether an experimental result is sufficiently (i.e., significantly) distinct from a random result; smaller *P*-values confer greater confidence that one has documented a truly significant result. The scientific convention for statistical significance is *P*-value • 0.05, or one chance in twenty. *See generally id.* at 272-78, 401-02, 415-20 (explaining the difference between causation and mere correlation, why regression analysis alone is inadequate to establish true causation, and the significance of the *P*-value).

88. *See* NETER ET AL., *supra* note 73, at 78-85.

89. The coefficient of determination (r^2) was calculated using MINITAB Statistical Software (Version 12.0). For instructions on calculating r^2 by hand, see MOORE, *supra* note 1, at 257-64.

90. *See supra* note 86.

91. The *P*-value was calculated using MINITAB Statistical Software (Version 12.0). *P*-value estimates could, however, be ascertained from a normal probability distribution table. *See infra* note 94 (describing the importance of the normal probability distribution). For reference purposes, compare these results with the hypothetical data depicted in graph a.) of Figure 1: the

Committee's conclusion, the essential message of formal regression analysis is not that chlorophyll *a* concentration and UKL elevation are uncorrelated, but that the available data are not capable of supporting *any* conclusion drawn from regression analysis. This discovery illustrates why an analytical technique that makes more judicious use of the data is needed. Accordingly, the next section of the Comment will introduce and perform just such an analysis.

C. *Taking a Closer Look at the FWS 2001 BiOp: the Mann-Whitney Two-Sample Test*

For the reasons discussed above, regression analysis was not a befitting test of the FWS 2001 BiOp. Closer inspection of the FWS 2001 BiOp does, however, suggest a superior alternative. Because the FWS 2001 BiOp prescribed specific, minimum monthly lake elevations,[92] it is possible to divide the chlorophyll *a* concentration data into two groups – those observed during years when UKL elevation satisfied the FWS 2001 BiOp requirement, and those from years when it did not – in a completely objective fashion. These two groups can then be compared directly, using the Mann-Whitney Two-Sample Test (Mann-Whitney Test).[93]

---

strength of association (r^2) between chlorophyll *a* and lake elevation in graph a.) is 99%, and the probability of the correlation being due to random variation is less than 0.1% (P-value < 0.001).

92. *See supra* note 61.

93. The Mann-Whitney Test is one of the most commonly used statistical procedures for comparing the central tendencies, or average values, of two experimental groups. It allows the analyst to determine whether one experimental group is significantly larger or smaller than the other. *See* JERROLD H. ZAR, BIOSTATISTICAL ANALYSIS 146-54 (4th ed. 1999); *see also* GOTTFRIED E. NOETHER, INTRODUCTION TO STATISTICS: THE NONPARAMETRIC WAY 103-113 (1991) (providing complete details on the use of the Mann-Whitney Test).

The Mann-Whitney Test is a nonparametric,[94] ranks-based procedure. It enumerates the number of times each data point in one experimental group (chlorophyll *a* concentration in years when UKL elevation did not meet the FWS 2001 BiOp recommendation) is larger, or "ranks higher," than any point in the second experimental group (chlorophyll a concentration in years when UKL elevation did satisfy the FWS 2001 BiOp), and vice versa.[95] The objective of the Mann-Whitney Test is, therefore, to determine whether the observations in the first group are, on average, larger or smaller than the observations in the second group.

---

94. There are two different kinds of two-sample tests: parametric and nonparametric. The Mann-Whitney Test is nonparametric, meaning that it does not rely upon a predefined probability distribution. A probability distribution reflects a trend within a population. Imagine, for example, that the FWS institutes a complete moratorium on UKL fishing in order to protect endangered suckers. FWS might be interested in determining whether the ban leads to an increase in the average size of adult suckers. Accordingly, it begins an aggressive sampling program and tracks the average size of adult suckers over the next 10 years. Assume also that adult suckers can be reasonably classified as "small," "medium," or "large." By the tenth year, FWS may determine that it is catching mostly medium-size suckers, with lesser numbers of small and large individuals. It is, at this point, difficult for FWS to conclude whether or not the moratorium has had the intended effect, because most natural populations of fishes – indeed, most populations of *any* species – have a *normal* probability distribution, relative to adult size. That is, individuals of "medium" size are almost always more common than individuals of very small, or very large size. Thus, FWS would need some means of determining whether their observed data are distributed in a manner that differs, to a biologically important degree, from the normal expectation.

Parametric statistical tests require *a priori* knowledge of the underlying probability distributions of one's experimental subjects. The "*t*-test" is the most commonly used parametric procedure for comparing the average values of two experimental populations, such as the average chlorophyll *a* concentrations in two different lake elevation groups. The *t*-test assumes a normal probability distribution (i.e., many observations within the midrange values and relatively few observations at either extreme), which is also known as the "bell curve." (There are many other probability distributions, but the normal distribution is the most common.) The *t*-test also assumes that the variability, or data scatter, within each of the two experimental groups is approximately equal; this second criterion is referred to as "homogeneity of variance." For example, if the observations in one experimental group were all of approximately equal value, while the observations in a second group spanned a large range of values, the assumption of homogeneity of variance would be violated, and the results of a parametric *t*-test would be suspect. *See generally* ZAR, *supra* note 93, at 65-85 (describing the theory and application of probability distributions and two-sample statistical tests); MOORE, *supra* note 1, at 226-32 (describing the critical characteristics of the normal probability distribution).

Unlike the parametric *t*-test, the nonparametric Mann-Whitney Test does not require *a priori* knowledge of the underlying probability distribution, nor does it assume that the data within each of the experimental groups are equally variable (i.e., homogeneity of variance). It is therefore more robust (i.e., less prone to error) than the *t*-test. This is particularly important when working with small datasets, for which the assumptions of a normal probability distribution and homogeneity of variance are difficult to verify. *See generally* ZAR, *supra* note 93, at 145-54 (explaining the Mann-Whitney Test and the basic theory of nonparametric statistics).

95. The Mann-Whitney Test ranking scores are expressed as "*U*" statistics. ZAR, *supra* note 93,at 146-48.

The Mann-Whitney Test makes better use of the empirical data than regression analysis in that it tests the FWS 2001 BiOp in a more acute manner. Rather than looking for a general, linear relationship between chlorophyll *a* concentration and lake elevation (i.e., regression analysis), the Mann-Whitney Test will only attempt to determine whether chlorophyll *a* concentration was significantly lower in years when UKL was at or above the elevation prescribed in the FWS 2001 BiOp than in years when UKL fell below that elevation. This test will preserve the overall logic of expecting higher lake elevations to benefit endangered suckers in UKL via improved water quality (i.e., lower algal densities, as reflected by chlorophyll *a* concentrations),[96] while simultaneously incorporating the actual FWS 2001 BiOp recommendations.[97] Before running the Mann-Whitney Test, however, it is expedient to first clarify precisely what weight of evidence will be required to prove that the FWS 2001 BiOp was either correct or incorrect, with regard to the prescribed UKL elevations.[98]

D. Quantifying Both Sides of the Argument: Power Analysis

As mentioned previously in the Comment, scientific peer review normally requires 95 percent certainty to claim a significant effect.[99] This means that scientists are, on average, willing to mistakenly claim a significant effect 5% of the time. Such a tautologous standard is both a curse and a blessing. For one, it preordains scientists to fail to detect many important, but as yet statistically insignificant effects.[100] On the other hand, an explicit significance criterion permits one to quantify the probability of failing to detect a significant effect, such as the effect of a specific lake elevation on chlorophyll *a* concentration, when it does, in fact, exist. This procedure, known as power analysis,[101] is discussed below.

Statistical power ("power") is the probability that a formal, quantitative statistical test, such as the Mann-Whitney Test, will produce a significant result.[102] Power is a function of three independent

96. *See supra* text accompanying notes 62-69.

97. The recommendations are listed *supra* note 61.

98. *Id.*

99. *See supra* note 49.

100. *See supra* note 1.

101. *See* Randall M. Peterman, *Statistical Power Analysis Can Improve Fisheries Research and Management*, 47 CAN. J. FISH. & AQUATIC SCI. 2 (1990); *see also* JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 1-17 (2d ed. 1988) (providing a complete introduction to the use of statistical power analysis).

102. Strictly speaking, power is the probability of correctly identifying a significant effect. *See* Catherine A. Toft & Patrick J. Shea, *Detecting Community-Wide Patterns: Estimating Power Strengthens Statistical Inference*, 122 AM. NATURALIST 618 (1983) (providing the conceptual basis of statistical power and relating its specific importance in endangered species management). Relative to the proposed Mann-Whitney Test, power is the probability of

parameters: the *a priori* defined criterion for claiming a significant effect (e.g., 95% certainty), sample size (the number of samples, or experimental observations in a dataset), and the "effect size" (the magnitude of the difference between experimental groups) that one is trying to detect.[103] Power is inversely proportional to the specified significance criterion; by trying to assure that they do not mistakenly claim a significant effect (i.e., demanding a high margin of certainty), one necessarily increases the likelihood of failing to detect an effect.[104] Sample size and effect size are directly proportional to power; larger sample sizes and larger effect sizes increase the likelihood of successfully detecting a significant difference among experimental groups.[105]

Power is important because it is a formal measure of the reliability of a statistical analysis.[106] Whenever a statistical test fails to detect a significant effect, but the associated power level is low, it is invalid to conclude that "no effect exists."[107] This Comment uses power analysis to determine the probability of having successfully detected a significant difference in chlorophyll *a* concentration between years when UKL elevation was at or above the level recommended in the FWS 2001 BiOp, and years when it was below that level, using only the nine empirical data points.[108] Power analysis can also be used to estimate the smallest effect size (e.g., reduction in chlorophyll *a* concentration) that could possibly be detected for a given power level and sample size,[109] or to determine how

---

detecting a significant difference between the two groups of chlorophyll *a* concentrations (i.e., years when UKL elevation was at or above 4141 feet and years when it was not). *See supra* text accompanying notes 92-97. Statistical power calculations are based on the same underlying probability distributions that are used in two-sample tests. *See supra* note 94.

103. Peterman, *supra* note 101, at 5.

104. *Id.*

105. *Id.*

106. *See generally* Peterman, *supra* note 101 (describing the use of power analysis as a formal indicator of the reliability of study designs and statistical tests). *See also* Toft & Shea, *supra* note 102, at 621-24 (providing additional examples of the use of power analysis as a formal measure of statistical reliability).

107. Peterman, *supra* note 101, at 2; *see also* Toft & Shea, *supra* note 102.

108. *See supra* note 70.

109. To estimate the smallest effect size that could be detected for a given confidence level, using only the existing chlorophyll *a* concentration and lake elevation data, it is first necessary to assign an appropriate "power level" (i.e., the degree of certainty required to conclude that chlorophyll *a* concentration is *not* significantly less when UKL elevation is at or above the FWS 2001 BiOp recommendation). In order to be just as confident that we will not miss a significant lake elevation effect as we are that we will not "jump the gun" and conclude there is an effect when, in fact, there is not, power must be set at 95%. At the 95% power level, we find that it would only be possible to detect a significant lake elevation effect if the average August chlorophyll *a* concentration in UKL *decreased by more than 90%*. This is a ridiculous proposition, because it means that chlorophyll *a* concentration must decline to virtually zero before we can hope to detect a significant lake elevation effect. Alternatively, if we are willing to lower our power level to 80%, we find that average August chlorophyll *a* concentration must be reduced by 62% before we could detect a significant lake elevation effect. *See generally*

many additional samples would be required to have a reasonable chance of detecting a particular effect size.[110] Combined with the results of the Mann-Whitney Test, power analysis will now be used to relate a more equitable assessment of the FWS 2001 BiOp than was employed in the NRC Committee's review.[111]

E. A More Equitable Pair of Results

In order to run the Mann-Whitney Test, the empirical chlorophyll *a* concentration and lake elevation data must first be segregated into two experimental groups. This division is illustrated in Figure 2, above. One data point is plotted for the August average chlorophyll *a* concentration in each of the nine years (1990-1998) that were available to the NRC Committee.[112] These data are divided in Figure 2 by the heavy, dashed vertical line, which reflects the FWS 2001 BiOp recommendation for 4141 feet of water in August;[113] there are five chlorophyll *a* concentration data points from years when average UKL elevation was below 4141 feet in August (75, 113, 115, 147, and 150 micrograms per liter), and four from years when it was at or above 4141 feet (75, 113, 155, and 197 micrograms per liter).

When the Mann-Whitney Test is performed for the nine years of UKL water quality data, the probability of chlorophyll *a* concentration being less when average August lake elevation equals or exceeds 4141 feet than when it is below 4141 feet turns out to be only 32%.[114] In other

---

Peterman, *supra* note 101 (describing the procedure for estimating detectable effect sizes). Power analysis calculations were performed with the Number Cruncher Statistical System (NCSS Statistical Software 2004).

110. Within an explicit, power analysis context, scientific convention tends to prefer 80% confidence (i.e., an 80% power level) that we will not fail to detect a significant effect, rather than 95%. Using the conventional standards of scientific research – a 95% significance criterion and 80% power level – to determine how many samples would be necessary to detect a given effect size, we find that, with a total of 18 samples (9 from years when UKL elevation was below 4141 feet, and 9 from years when it was at or above 4141 feet), it would be possible to detect a 50% reduction in chlorophyll *a* concentration. Detecting a 20% reduction, however, would require approximately 43 annual samples from each of the experimental groups, or 86 samples total. *See generally* Peterman, *supra* note 101 (describing the procedure for determining the minimum sample size necessary to detect a particular effect size).

111. For additional examples of the use of power analysis in ESA management, see Barbara L. Taylor & Tim Gerrodette, *The Uses of Statistical Power in Conservation Biology: The Vaquita and Northern Spotted Owl*, 7 CONSERVATION BIOLOGY 489 (1993). *See also* Randall M. Peterman, *Application of Statistical Power Analysis to the Oregon Coho Salmon (Oncorhynchus kisutch) Problem*, 46 CAN. J. FISH. & AQUATIC SCI. 1183 (1989).

112. FINAL REPORT, *supra* note 5, at 110-13.

113. FWS 2001 BIOP, *supra* note 22, at § 3(2) 143.

114. If a parametric *t*-test is used instead of the nonparametric Mann-Whitney Test, the corresponding probability of a significant effect is still 32%. *See supra* note 94 (describing the differences between the nonparametric Mann-Whitney Test and the alternative procedure – the parametric *t*-test).

words, science is *32% certain* that maintaining 4141 feet of water in August would improve UKL water quality by abating chlorophyll *a* concentration (i.e., algal growth). Clearly, there is little empirical justification for implementing the FWS 2001 BiOp recommendation that 4141 feet of water be left in UKL during the month of August. However, there is not yet adequate evidence to prove that the USFWS 2001 BiOp was wrong. Such a complete result requires power analysis.

Power analysis is used here to calculate the level of certainty with which one may conclude that chlorophyll *a* concentration is not significantly less when UKL elevation is at or above 4141 feet in August, than when it is below 4141 feet. Bound to a 95% certainty criterion, power analysis of the Mann-Whitney Test result (32% certainty of a significant difference between the two experimental groups) shows that science is only *2% certain* that keeping at least 4141 feet of water in UKL in August will *not* significantly reduce chlorophyll *a* concentrations.[115] At this point, science appears to have reached an impasse; the existing data provide 32% certainty that lake elevations at or above 4141 feet will reduce chlorophyll *a* concentrations, and 2% certainty that they will not. Obviously, science is not yet confident of either possibility, and, in this particular situation, greater certainty will only be achieved with the benefit of additional data. This has, however, been a valuable exercise, because it illustrates a decisive limitation of the scientific process: if science fails to prove that a given theory is true, such as an inverse relationship between lake elevation and algal growth (as indicated by chlorophyll *a* concentration), *it does not immediately follow that the theory is false.*

Comparing the Mann-Whitney Test and power analysis results directly provides one final, compelling piece of information. To conclude that August lake elevations below 4141 feet do not promote algal blooms in UKL, based solely upon the nine available data points, is to assume, quite literally, that mistakenly implementing the FWS 2001 BiOp lake elevation recommendations (i.e., maintaining the recommended monthly minimum lake elevations when they do not truly mitigate algal blooms) is approximately *20 times more reprehensible* than the opposite mistake – failing to provide services that are, in fact, necessary to protect endangered suckers.[116] The next Part of the Comment explains why such

115. If a *t*-test is used in place of the Mann-Whitney test, the corresponding probability of failing to detect a significant difference in chlorophyll *a* concentration, when it does exist (i.e., the power level), is *11%*. *See supra* note 94.

116. The "20 times" estimate is the ratio of the probabilities of the two potentially mistaken conclusions. It is calculated by subtracting each of the percent certainty values from 100% (i.e., absolute certainty), and dividing the probability of failing to notice a significant lake elevation effect (100% – 2% = 98%) by the probability of mistakenly concluding that lake elevation has a significant effect on chlorophyll *a* when, in fact, it does not (100% – 95% = 5%): 98 ÷ 5 = 19.6.

a colossal supposition is not only compatible with, but central to the practice of sound science. Meanwhile, Part I closes with an earnest caution: while this Comment has, for heuristic purposes, focused exclusively upon the NRC Committee's review of the FWS 2001 BiOp monthly lake elevation recommendations, the analytical process demonstrated herein is applicable to the entire content of the FWS 2001 BiOp,[117] as well as the National Marine Fisheries Service's 2001 biological opinion on threatened coho salmon in the lower Klamath River.[118] In fact, these considerations are germane to all scientific research.[119] Thus, the Comment's relevance extends far beyond the upper Klamath Basin.

II. THE LIMITS OF STATISTICAL INFERENCE: HOW SCIENTISTS DEAL WITH UNCERTAINTY

This Part of the Comment examines why the NRC Committee was so dismissive of the FWS 2001 BiOp recommendation for 4141 feet of water in August, and why other scientific, peer-review authorities are likely to adopt similarly unbalanced positions in the future. Although the NRC Committee did not employ formal statistical tests, it did conform to the essential logic of statistical inference.[120] Statistical inference is the systematic decision-making process that scientists use to draw conclusions from experimental data.[121] It dictates the necessary and sufficient conditions for establishing a causal relationship, or significant effect.[122] Such a methodical formula is indispensable to the scientific community because, as this Comment has already demonstrated, scientific research does not always lead to a definitive conclusion. Whenever multiple, but equally sound interpretations of equivocal data are possible, scientists use

---

See Toft & Shea, *supra* note 102, at 620 (explaining how this calculation is made, and illustrating its importance in endangered species management).

117. *See generally* FINAL REPORT, *supra* note 5, at 214-49 (detailing the NRC Committee's complete review of the FWS 2001 BiOp, including multiple water quality, physical habitat degradation, and physical entrainment concerns).

118. *See* NATIONAL MARINE FISHERIES SERVICE, BIOLOGICAL OPINION: ONGOING KLAMATH PROJECT OPERATIONS (April 6, 2001), *available at* http://www.usbr.gov/mp/kbao/esa/38_cohobo_4_6-01.pdf. *See also* FINAL REPORT, *supra* note 5, at 287-310 (detailing the NRC Committee's complete review of the NMFS 2001 Biological Opinion, including recommended instream flows for the lower Klamath River and juvenile coho habitat requirements).

119. *See generally* THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS (3d ed. 1996) (describing the historical growth and evolution of scientific paradigms). *See also* KARL R. POPPER, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE (2d ed. 1965) (providing a more detailed account of inferential logic and hypothesis testing).

120. *See* FINAL REPORT, *supra* note 5, at 33-37 (describing the logic that was used by the NRC Committee to assess uncertainties within the FWS 2001 BiOp).

121. *See* MOORE, *supra* note 1, at 330, 377-79.

122. *See generally id.* at 397-403 (describing the basic theory of statistical inference testing).

statistical inference to ensure an unbiased outcome.[123] Statistical inference is not, however, an invention for divining "truth." Rather, the immediate ambition of science, realized via statistical inference, is unconditional objectivity.[124] Accordingly, the Comment continues with an introduction to statistical inference, beginning with the formulation of scientific hypothesis tests.

A. *Formulating Hypothesis Tests*

Scientists structure their research in terms of hypothesis tests.[125] Every scientific study involves a null hypothesis (notated as "H_O") and an alternative hypothesis (notated as "H_A").[126] The null hypothesis states that there has been "no effect" or that there is "no difference" between two experimental units.[127] The alternative hypothesis is the assertion that the scientist is trying to prove; it postulates that some significant effect or change has occurred.[128] In a test of statistical inference, the null hypothesis always receives the benefit of doubt,[129] while the alternative hypothesis assumes the burden of proof.[130] For example, the NRC Committee's review of the FWS 2001 BiOp can be used to relate the anatomy of a hypothesis test: the FWS 2001 BiOp claim that 4141 feet of water in August would reduce chlorophyll *a* concentration was the alternative hypothesis, while the competing claim, that 4141 feet of water would not reduce chlorophyll *a* concentration, was the null hypothesis.

Once the null and alternative hypotheses have been defined, the scientist's task is to gather evidence in support of the alternative hypothesis, and determine whether this evidence is sufficient to reject the null hypothesis with a high degree of certainty (e.g., 95% certainty).[131] As the instrument of that determination, statistical inference imbues the process with impartiality and consistency. Statistical inference does, however, confer a categorical advantage to the null hypothesis; anytime the scientist fails to accumulate sufficient proof of the alternative, she is obligated, under the tenets of peer review, to defer to the null hypothesis.[132] This explains why the NRC Committee, as a peer-review jury, did not extend its assessment of the FWS 2001 BiOp's prescribed

---

123. *See generally* MOORE, *supra* note 1, at 331-33, 377-81 (relating the use of statistical inference testing to ensure unbiased results).

124. *See id.*

125. *Id.* at 400-02.

126. *Id.* at 400.

127. *Id.*

128. *Id.*

129. *Id.*

130. *Id.*

131. *See* MOORE, *supra* note 1, at 400-402.

132. *Id.*

minimum monthly UKL elevations beyond a simple dismissal of the alternative hypothesis (i.e., that 4141 feet or more of water in August would abate algal growth).[133] It also alludes to a fundamental disparity between the objectives and resources of unadulterated, peer-reviewed science and conservation applications, such as endangered species management. In order to address this disparity, the next sections of this Comment examine why science tends to favor the null hypothesis, and how a more balanced approach might be realized, relative to ESA decisions.

B. Type I versus Type II Error

At the conclusion of a hypothesis test, one of four outcomes is possible (Figure 3).[134] If the alternative hypothesis is true, and a test of statistical inference is concordant, the scientist will correctly conclude that she has documented a significant effect.[135] Similarly, if the null hypothesis is true, and a test of statistical inference fails to validate the alternative hypothesis, the scientist will make the correct conclusion.[136] If, on the other hand, the null hypothesis is true, but the scientist mistakenly rejects it, a false-positive, or Type I error, occurs (e.g., erroneously concluding that 4141 feet of water in August is necessary to control algal growth and protect the endangered suckers in UKL).[137] Or, if the alternative hypothesis is true, but the null hypothesis is mistakenly accepted, a false-negative, or Type II error, occurs (e.g., concluding that 4141 feet of water in August will not benefit endangered suckers in UKL by reducing algal concentrations when, in fact, they will).[138]

---

133. Because the FWS 2001 BiOp did not support its minimum monthly UKL elevation recommendations with sufficient empirical proof, the NRC Committee accepted the null hypothesis (i.e., that 4141 feet of water would not reduce algal growth) by default. *See supra* note 74.

134. *See* MOORE, *supra* note 1, at 422-26.

135. *Id.*

136. *Id.*

137. *Id.*

138. *Id.*

True Condition (unknown)

	H_O (4141 feet of water *will not* reduce algal density)	H_A (4141 feet of water *will* reduce algal density)
Accept H_O	Correct conclusion	Type II error
Accept H_A	Type I error	Correct conclusion

Figure 3. The four possible outcomes of a hypothesis test, relative to the effects of 4141 feet of water on algal density.

By assigning the burden of proof to the alternative hypothesis (e.g., requiring the FWS to prove that its monthly lake elevation recommendations will mitigate algal growth, rather than requiring the DOI to prove that KP operations are not facilitating algal growth), statistical inference ensures that Type I errors will be committed less often than Type II errors.[139] Science is normally tolerant of these "missed opportunities," or Type II errors, for several reasons.[140] First of all, failing to recognize a significant treatment effect is not especially problematic when an entire community of one's peers is waiting to perform similar

139. *See* MOORE, *supra* note 1, at 414-16, 422-26.
140. *See id.*; *see also* POPPER, *supra* note 119, at 215-48.

tests. Given sufficient means, scientists trust that the legitimacy of a correct, but as yet unproven hypothesis will be demonstrated, and so they tend to envisage Type II error as more of a temporary delay than a legitimate failure.[141] From the scientist's perspective, a more reprehensible mistake is the reporting of a false discovery, or Type I error,[142] as the following heuristic illustrates.

Scientific research is similar to a branching network of roads, where each hypothesis test represents a junction within the network.[143] At any given junction, the scientist must decide whether to continue down one of the branches (i.e., accept an alternative hypothesis), or to remain at the junction (i.e., accept the null hypothesis), awaiting greater certainty of the path that should be followed.[144] In theory, it costs the scientist little to bide their time at a junction, while performing additional tests. To reach a premature (and ultimately mistaken) conclusion and proceed down the wrong path is, however, a potentially significant problem. Because science is such an iterative process, with each successive generation of research questions depending heavily upon the principles established by its predecessors,[145] the *post hoc* discovery of a Type I error forces the scientist to retrace her progress through the network, and possibly discard whatever results have been compiled since turning onto the wrong road. A Type I error is also likely to cost the scientist some part of her professional credibility, particularly when her peers have already conducted research that depended upon the validity of her mistaken conclusion, and must therefore backtrack as well. Thus, the rationale behind science's conservative outlook stems largely from its sequential, step-by-step mechanism.[146]

The conservative ethos of science and statistical inference is also analogous to the "innocent until proven guilty" maxim of criminal law,[147] and the "first do no harm" axiom of the medical profession.[148] All three assume that a Type I error (i.e., mistakenly claiming that a significant effect, change, or event has occurred), whether it be the unwarranted cessation of an irrigation project, the misguided conviction of an innocent

141. *See* POPPER, *supra* note 119, at 215-48.

142. *See* MOORE, *supra* note 1, at 415-16.

143. *See generally* John R. Platt, *Strong Inference*, 146 SCIENCE 347 (1964) (describing the "branching" nature of scientific research and the importance of formal, inferential logic).

144. *See id.*

145. *See id.*; *see also* POPPER, *supra* note 119, at 238-48.

146. *See* POPPER, *supra* note 119, at 215-48.

147. "The principle that there is a presumption of innocence in favor of the accused is the undoubted law, axiomatic and elementary, and its enforcement lies at the foundation of the administration of our criminal law." Coffin v. United States, 156 U.S. 432, 453 (1895).

148. The origin of the "do no harm" standard is generally attributed to the Greek philosopher Hippocrates: "As to diseases, make a habit of two things – to help, or at least to do no harm." MAURICE B. STRAUSS, FAMILIAR MEDICAL QUOTATIONS 625 (1968).

man, or the accidental poisoning of a patient, is a more grave proposition than the opposite, Type II error (i.e., failing to notice that a significant effect, change, or event has occurred). This conservative paradigm becomes ill-founded, however, whenever the consequences of a Type I error rival or exceed those of a Type II error.[149] For example, a terminally-ill patient may decide that it is prudent to risk the adverse side effects of an experimental, but potentially lifesaving treatment.[150] Likewise, the survival of an endangered species may depend upon the willingness of scientists, lawyers, and politicians to proactively institute plausible, but currently unproven protective remedies.[151] As statistician David S. Moore points out, there are scenarios in which the indiscriminate postulation that Type II error is preferable to Type I error is a fallacy, and a more balanced approach becomes justified:

> Tests of significance concentrate attention on H_0, the null hypothesis. If a decision is called for, however, there is no reason to single out H_0. There are simply two alternatives, and we must accept one and reject the other. It is convenient to call the two alternatives H_0 and H_a, but H_0 no longer has the special status... that it had in tests of significance.... There is no reason to put the burden of proof on the [alternative hypothesis]... unless we have strong evidence against it. *It is equally sensible to put the burden of proof on the [null hypothesis]...*"[152]

C. *Balancing the Costs of Uncertainty in Science and Policy*

To concurrently employ both "purely scientific" and "ESA compliant" standards in the face of substantial uncertainty will, at times, be impossible, because the two doctrines have been fashioned to prevent different types of errors; scientific peer review is most concerned with preventing Type I errors, while the proactive ESA seeks to prevent Type II error.[153] This is why the FWS and the NRC Committee, upon analyzing the same data, reached opposite conclusions. As the guardian of the endangered Lost River and shortnose suckers, the FWS was charged with preventing a Type II error (i.e., failing to institute procedures that would

149. *See* MOORE, *supra* note 1, at 422-26.

150. In the case of a terminally-ill patient, the administration of an ineffective treatment, and consequent exposure to unknown side effects, would constitute a Type I error. By contrast, the failure to administer an effective treatment would constitute a Type II error.

151. Relative to the recommended FWS 2001 BiOp minimum monthly lake elevations, a Type I error would occur if water were withheld from KP irrigators, but higher lake elevations did not actually reduce algal concentrations. A Type II error would occur if the action agencies failed to recognize a significant (i.e., inverse) effect of lake elevation on algal concentration, therefore missing the opportunity to institute procedures that would have benefited endangered suckers in UKL.

152. MOORE, *supra* note 1, at 423 (emphasis added).

153. *See supra* note 46 (citing the proactive language of the ESA, relative to agency actions).

benefit endangered suckers in UKL), while the NRC Committee was obligated to prevent a Type I error (i.e., mistakenly deciding that the FWS 2001 BiOp prescriptions for minimum monthly lake elevations were truly necessary to protect endangered suckers).[154] Given these contradictory imperatives, society can approach the situation in one of several ways. It can apply the scientific, statistical inference standard, which the NRC Committee used,[155] knowing this will place the burden of proof on the endangered suckers. Alternatively, it can apply the precautionary standard the FWS used,[156] thereby assigning the burden of proof to the KP. Or, society can choose to recognize that this is not an issue that scientific authorities alone can resolve.[157]

Critics of the Interim Report contend that the NRC Committee should have applied the "precautionary principle."[158] The precautionary principle makes reversing the burden of proof a formal policy matter, by requiring agencies to demonstrate that a proposed action will not harm

---

154. The NRC Committee acknowledged the fundamental difference between its own mission, as a scientific, peer-review authority, and the proactive mission of the FWS:

> The committee, in drawing conclusions for its interim report, was bound by its charge to evaluate and comment on the scientific strength of evidence underlying various proposals. Its charge kept it from weighing economic concerns or weighing the advisability of minimizing risk by using professional judgment in place of scientific evidence to support particular recommendations. . . . [A]gencies charged with ESA responsibilities can be expected to use professional judgment when no scientifically supportable basis is available for a decision, or where they judge the scientific support to be inadequate. Thus, the agencies may recommend practices for which the committee would find virtually no direct scientific support. The committee acknowledges the necessity of this practice in many situations where information is inadequate for development of scientifically rigorous decisions.

FINAL REPORT, *supra* note 5, at 35.

155. *Id.*

156. *Id.* at 34.

157. In a previous report on endangered species management, the National Research Council's Committee on Scientific Issues in the Endangered Species Act stated: "Although the language of the ESA suggests that the standards for making decisions about listing, jeopardy, etc., are to be purely scientific, analyses of ESA implementation show clearly that tradeoffs among conflicting objectives must be made in almost every instance." NATIONAL RESEARCH COUNCIL, SCIENCE AND THE ENDANGERED SPECIES ACT 134 (1995) (internal citations omitted).

158. Although multiple interpretations of the precautionary principle have been instituted in various international treaties, the basic concept is typically referenced to the Rio Declaration of the United Nations Conference on Environment and Development: "In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation." Rio Declaration on Environment and Development, United Nations Conference on Environment and Development, Annex 1, U.N. Doc. A/CONF.151/26 (1992), *available at* http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm; *see also* Kenneth R. Foster et al., *Science and the Precautionary Principle*, 288 SCIENCE 979 (2000) (discussing recent applications of the precautionary principle and guidelines for its enforcement).

ESA listed species, rather than requiring ESA advocates to prove that said action will, in fact, injure listed species.[159] Under the precautionary principle, the FWS 2001 BiOp would have received the benefit of doubt, while the DOI would have been required to prove that KP operations were not harming the endangered suckers. In the Final Report, the NRC Committee defended its judgment by pointing out that "whether to apply the precautionary principle is a policy decision and as such is outside the present committee's scope of work, which pertains to 'whether the biological opinions are consistent with the available scientific information.'"[160] Neither of these dispositions are particularly objective, as they both stem from institutional assumptions about the gravity of Type I and II errors; the precautionary principle is a tool for preventing Type II errors, while the conservatism of statistical inference is a means of preventing Type I errors. However, when science provides quantitative measures of both Type I and II error probabilities, it becomes possible to make balanced decisions, instead of relying upon such tautologous assumptions.[161]

The chlorophyll *a* concentration and UKL elevation test results reported in Part I.E of the Comment can be used to quantify the corresponding Type I and Type II error probabilities. Together, the Mann-Whitney Test and power analysis showed that there is a 32% probability of the alternative hypothesis being correct (i.e., that maintaining August UKL elevations at or above 4141 feet will reduce chlorophyll *a* concentration), and a 2% probability of the null hypothesis being correct (i.e., that 4141 feet of water in August will not reduce chlorophyll *a* concentration). Thus, the likelihood of committing a Type I error, without the benefit of additional data, is 68%,[162] while the

159. The National Research Council has previously acknowledged the special nature of ESA decisions, when implemented under the precautionary principle:

> If the burden of proof were to show that an action would not harm a species rather than to show that it would harm a species, increased protection would result. The importance of shifting the burden of proof this way has been widely recognized, especially in the context of marine conservation issues, and is known as the "precautionary principle."

NATIONAL RESEARCH COUNCIL, *supra* note 157, at 133. An example of the successful application of the precautionary principle is provided by the migratory Snake River sockeye salmon (*Oncorhynchus nerka*). In 1991, the migratory Snake River (Idaho) sockeye population was listed as endangered under the ESA, despite the fact that scientists had not yet conclusively proven its imperilment. In making this determination, the NMFS explicitly considered the costs of failing to institute protective status, in accordance with the precautionary principle. *See* Robin S. Waples, *Evolutionarily Significant Units and the Conservation of Biological Diversity under the Endangered Species Act*, *in* EVOLUTION AND THE AQUATIC ECOSYSTEM: DEFINING UNIQUE UNITS IN POPULATION CONSERVATION 8 (Jennifer L. Nielson ed., 1995).

160. FINAL REPORT, *supra* note 5, at 315.

161. *See* MOORE, *supra* note 1, at 422-26.

162. The Type I error probability is obtained by subtracting the corresponding percent certainty value (i.e., 32% certainty of the alternative hypothesis) from 100%. *See* MOORE, *supra* note 1, at 425.

likelihood of committing a Type II error is 98%.[163] Armed with these figures, it is now feasible to relinquish the conservative, Type I error prevention ethic of statistical inference, and make the transition to a more equitable decision making process, in which the probabilities and costs of both Type I and II errors are explicitly considered.[164] Such a balanced, inclusive approach was advanced by Doremus and Tarlock: "Ultimately, science alone cannot tell us how to allocate the limited water resources of the Klamath Basin. Neither can a simple-minded appeal to caution. Society must decide how cautious it should be, and at what cost."[165] This approach is unlikely to be realized, however, if Congress passes the Data Quality Act, because the Data Quality Act would make the conservative norms of peer-reviewed science and statistical inference a legally-binding mandate. The next Part of the Comment is therefore dedicated to clarifying a number of the Data Quality Act's critical shortcomings.

III. THE "ENDANGERED SPECIES DATA QUALITY ACT OF 2004"

At a glance, the Data Quality Act appears to be a reasonable petition for infusing the ESA process with more rigorous scientific expertise. It would require the Secretary to "give greater weight to scientific or commercial data that is empirical or has been field-tested or peer-reviewed" anytime multiple or conflicting datasets are evaluated under the "best scientific and commercial data available" standard.[166]

163. The Type II error probability is obtained by subtracting the corresponding percent certainty value (i.e., 2% certainty of the null hypothesis, when 95% certainty is required to accept the alternative hypothesis) from 100%. *See id.*

164. Balancing Type I and II error probabilities is not necessarily a simple question of avoiding action associated with the lower probability; although the probability of committing a Type II error (i.e., 98% probability of failing to detect a significant lake elevation effect on algal density) is greater than the probability of committing a Type I error (i.e., 68% probability of mistakenly concluding that higher lake elevations abate algal growth), a truly balanced decision will include economic and sociological factors. Formal "decision analysis" methods for integrating Type I and II error probabilities with additional sources of information, such as financial commitments and limitations, are available. These methods provide a systematic alternative to the unilateral, Type I and II error prevention mechanisms of the precautionary principle and scientific inference testing. Decision analysis techniques are, however, beyond the scope of this Comment. For a brief introduction to decision analysis, see NATIONAL RESEARCH COUNCIL, *supra* note 157, at 136. *See also* ROBERT T. CLEMEN, MAKING HARD DECISIONS: AN INTRODUCTION TO DECISION ANALYSIS (2d ed. 1997) (providing complete instruction on the application of decision analysis).

165. Doremus & Tarlock, *supra* note 4, at 343.

166. *See* Data Quality Act, S. 2009, 108th Cong. § 2(a) (2004). Included under the "best scientific and commercial data available" blanket would be the review of petitions for listing and de-listing; 16 U.S.C. § 1533(b)(1)(A) (2000); critical habitat designations and the option to overrule cost-benefit exclusions of critical habitat on threat of imminent extinction; § 1533(b)(2); verification of the presence of endangered species within the area of a proposed agency action and the pursuant obligation to complete a biological assessment; § 1536(c)(1); agency consultations; § 1536(a)(2); jeopardy opinions and reasonable and prudent alternatives; §

With specific regard to listing (but not de-listing) decisions, the Data Quality Act would require that the Secretary "establish [explicit] criteria that must be met for scientific and commercial data to be used as the basis of a determination... that a species is... endangered... or threatened"[167] and that all future listings be "supported by data obtained by observation of the species in the field."[168] And it would institute mandatory peer review of all ESA listings, de-listings, recovery plans, and jeopardy opinions.[169]

Dig just a little beneath the surface, however, and the Data Quality Act loses its benign character. Two points are particularly insidious: the addition of explicit listing criteria,[170] and the peer-review requirements.[171] Consistent with scientific custom, explicit listing criteria are likely to demand 95% certainty of "threatened" or "endangered" status (which is the alternative hypothesis, because it implies a significant departure from the status quo), without ensuring a similarly high power level (i.e., the probability of proving that a species is truly threatened or endangered). Such a biased policy could prove crippling to listing advocates because, as this Comment has pointed out, statistical power is a function of sample size and the magnitude of the effect one is attempting to detect.[172] As a specific example, the Comment determined that 18 annual samples – twice as many as the NRC Committee had at its disposal – would be necessary to have a high probability of successfully detecting a 50% reduction in chlorophyll *a* (i.e., 80% power).[173] This result begs the importance, relative to explicit listing criteria, of designating equally explicit criteria for the amount of effort that must be dedicated to assembling an adequate dataset. In natural settings, where environmental variability tends to be high, it will often times be impossible to detect a significant effect (i.e., the threatened or endangered status of a species) without some minimum number of samples.[174] Therefore, requiring that explicit criteria be satisfied prior to ESA listing categorically disregards

---

1536(a)(2), (b)(3)(A); and the withdrawal of a permanent exemption for agency action, pursuant to the discovery of previously unnoticed endangered species; § 1536(h)(2)(B)(i).

167. S. 2009, 108th Cong. § 2(b)(9).

168. § 2(b)(10)(A). Field data may be submitted by private landowners, as well as agency personnel and third-party scientists; § 2(b)(10)(B). Note that, for the purposes of this Comment, we have elected not to deal with the appointment of peer reviewers, the solicitation of state information, or the invitation for greater public participation in the consultation process. § 3(j)(2)-(3), 4(a)-(b). These issues are not about "sound science," save, perhaps, for the fact that, in many instances, private landowners may not be qualified to identify or enumerate species of concern.

169. § 3(j)(1)(A)(i)-(iv).

170. § 2(b)(9).

171. § 3(j)(1)(A)(i)-(iv).

172. *See supra* text accompanying notes 102-105.

173. *See supra* note 110.

174. *Id.*

the probability and consequences of a Type II error (i.e., failing to list a threatened or endangered species), unless there is also an equally forthright guarantee that an adequate sampling effort, which has been identified via non-arbitrary methods (i.e., power analysis), will be devoted to satisfying those criteria.

Moreover, if the NRC Committee's peer-review analysis of the FWS 2001 BiOp becomes the model for subsequent ESA reviews,[175] then the primary effect of the Data Quality Act's peer-review requirements would be to entrench the Type I error prevention paradigm in all ESA recovery plans and jeopardy opinions, as well as listing decisions. And, unless the appointed, independent reviewers of ESA actions are willing to acknowledge the exceptional nature of those actions, and deliberately choose to incorporate both Type I and II error probabilities in their reviews, then enactment of the Data Quality Act would amount to surrendering the precautionary spirit of the ESA. After all, the conventional manifestation of scientific peer review does not consider Type II error probability.

Scientists appreciate the confusion that qualitative ESA standards, such as "the best scientific and commercial data available"[176] and "substantial scientific or commercial information,"[177] have caused for the action agencies, as well as the courts. Before the 109th Congress decides the fate of the Data Quality Act, however, its members should be prepared to wrestle with a difficult, and seemingly overlooked, question: is the purpose of the ESA simply to promote "sound science," in which case it is no longer unique from other scientific research initiatives, or is it

---

175. Several lines of evidence do suggest that the review process instituted under the Data Quality Act would be modeled after the Interim Report. For example, the Data Quality Act requires the Secretary of Interior to "solicit recommendations from the National Academy of Sciences and develop and maintain a list of qualified reviewers." S. 2009, 108th Cong. § 3(j)(2) (2004). In general, the Interim Report has been hailed by multiple agencies and persons as an exemplary instance of independent peer review. For example, Assistant Secretary of the Interior for Fish, Wildlife and Parks Craig Manson told the House Resources Committee in 2002:

> Before I discuss the specific provisions of the bill, I want to acknowledge that addressing these issues in any context is not an easy task, and I would like to commend the [NRC] Committee for its efforts in this regard. . . . In this respect, [the Data Quality Act] requires that an independent review of science be carried out by "qualified individuals," as determined by the National Academy of Science (NAS) standards. *The Department has had significant experience with the NAS review process, and is comfortable that this provision will help ensure a truly independent scientific review process.*

Testimony of Craig Manson Before the House Resources Committee (emphasis added), *available at* http://laws.fws.gov/testimon/2002/science.html (last visited February 20, 2005).

176. *See supra* note 166 (listing ESA actions that are subject to the "best scientific and commercial data available" standard).

177. ESA actions subjected to the "substantial scientific or commercial information" standard include the petitions for listing or de-listing and the review of petition for critical habitat modifications. 16 U.S.C. §1533(b)(3)(A), (b)(3)(D)(i) (2000).

to ensure the future of America's rich biological diversity? Perhaps the real contest is not to characterize sound science, but to determine whether it is wise to canonize science as the preeminent arbiter of ESA decisions. The scientist is no more qualified to answer this question than any other citizen. But it is worth noting that the National Research Council, whose opinions and activities are intimately linked to the current management of endangered fishes in UKL,[178] as well as Congress' interest in the Data Quality Act, proffered this poignant advice seven years prior to the publication of the Interim Report:

> In the usual procedures for formulating scientific tests of hypotheses, it is customary to... limit the probability of falsely rejecting [the] null hypothesis to a known level (often, 0.05) while permitting much larger probabilities of falsely accepting the null hypothesis. We are concerned that when such... procedures are followed in ESA decision making, they will too often place the burden of proof (for demonstrating a significant effect) on those who want to institute some protective actions (usually the FWS or petitioners for listing of a species), without taking into account the practical consequences of falsely concluding that no effect is occurring. *This could lead to a systematic bias against species that are candidates for listing or for listed species in need of protective actions.*[179]

CONCLUSION

While statistical inference does ensure a high degree of consistency in scientific research, it is ultimately an imperfect method, because it discriminates against those parties that must shoulder the burden of Type II error (i.e., failing to detect a truly significant effect). In ESA affairs, where the price of Type II error may be extinction, this inequity will tend to be particularly acute. Politicians, legal professionals, and natural resource managers can liberate themselves from the confines of this conservative, "Type I error world," without losing the benefit of science's technical expertise, in one of two ways: they can consciously choose to reverse the assumption that a Type II error is preferable to a Type I error, and adopt an endangered species philosophy that is more consistent with the precautionary principle. Or, they can require weight-of-evidence standards that explicitly address the probabilities and consequences of both Type I and II error, and relinquish the 95% certainty touchstone in favor of a more balanced approach. Neither of these strategies would be likely to constitute "arbitrary or capricious"

178. *See generally* Doremus & Tarlock, *supra* note 4, at 324-336 (describing the chain of events that followed the release of the Interim Report and that ultimately led to the Data Quality Act).

179. NATIONAL RESEARCH COUNCIL, *supra* note 157, at 168 (emphasis added).

behavior.[180] However, the purpose of this Comment is not to prescribe an appropriate course of action, but to elucidate the process of statistical inference testing, and make "sound science" more transparent.

The authors also wish to emphasize that the Comment's criticisms of the Interim and Final Reports do not implicate the professional expertise of the NRC Committee members. Indeed, the collective recommendations of the Final Report would seem to represent the Lost River and shortnose suckers' best chance for recovery.[181] Aside from the critique of regression analysis (relative to the chlorophyll *a* concentration and UKL elevation data), the concerns underpinning this Comment have only to do with the NRC Committee's apparent apathy towards audiences that were not likely to appreciate the technical suppositions underlying its generalized results.

Given that the authors, oracles, and agents of environmental policy operate in largely independent capacities, a limited measure of miscommunication amongst the three is understandable, and perhaps inevitable. But so long as scientists fail to recognize that their norms are not well understood by their interdisciplinary colleagues, society can never hope to achieve the full measure of congressional intent that is codified in environmental statutes. This Comment has attempted to identify and explain some of science's most critical limitations and assumptions, in the hope that such monumental institutions as the Endangered Species Act may not fall victim to the whims of obtuse parties.

---

180. Such behavior is prohibited under the Administrative Procedure Act, 5 U.S.C. § 706(2)(A) (2000).

181. *See generally* FINAL REPORT, *supra* note 5, at 214-249, 346-349 (providing a complete list of the challenges facing endangered sucker recovery efforts).