Polar Bear Population Forecasts: A Public-Policy Forecasting Audit

J. Scott Armstrong and Kesten C. Green and Willie Soon

The Wharton School, University of Pennsylvania, Business and Economic Forecasting Unit, Monash University, Harvard-Smithsonian Center for Astrophysics

15. December 2007

Polar Bear Population Forecasts: A Public-Policy Forecasting Audit

Working Paper Version27: Dec 15, 2007

J. Scott Armstrong, The Wharton School, University of Pennsylvania

747 Huntsman, Philadelphia, PA 19104, USA.
armstrong@wharton.upenn.edu; jscottarmstrong.com; T 610-622-6480

Kesten C. Green, Business and Economic Forecasting, Monash University, Vic 3800, Australia.
Contact: PO Box 10800, Wellington 6143, New Zealand.
kesten@kestencgreen.com; T +64 4 976 3245; F +64 4 976 3250

Willie Soon, Harvard-Smithsonian Center for Astrophysics, Cambridge MA 02138, USA.
wsoon@cfa.harvard.edu

Abstract

The extinction of polar bears by the end of the 21st century has been predicted and calls have been made to list them as a threatened species under the U.S. Endangered Species Act. The decision on whether or not to list rests upon forecasts of what will happen to the bears over the 21st Century.

Scientific research on forecasting, conducted since the 1930s, has led to an extensive set of principles—evidence-based procedures—that describe which methods are appropriate under given conditions. The principles of forecasting have been published and are easily available. We assessed polar bear population forecasts in light of these scientific principles.

Much research has been published on forecasting polar bear populations. Using an Internet search, we located roughly 1,000 such papers. None of them made reference to the scientific literature on forecasting.

We examined references in the nine unpublished government reports that were prepared "…to Support U.S. Fish and Wildlife Service Polar Bear Listing Decision." The papers did not include references to works on scientific forecasting methodology.

Of the nine papers written to support the listing, we judged two to be the most relevant to the decision: Amstrup, Marcot and Douglas et al. (2007), which we refer to as AMD, and Hunter et al. (2007), which we refer to as H6 to represent the six authors. AMD's forecasts were the product of a complex causal chain. For the first link in the chain, AMD assumed that General Circulation Models (GCMs) are valid. However, the GCM models are not valid as a forecasting method and are not reliable for forecasting at a regional level as being considered by AMD and H6, thus breaking the chain. Nevertheless, we audited their conditional forecasts of what would happen to the polar bear population *assuming* that the extent of summer sea ice will decrease substantially in the coming decades.

AMD could not be rated against 26 relevant principles because the paper did not contain enough information. In all, AMD violated 73 of the 90 forecasting principles we were able to rate. They used two un-validated methods and relied on only one polar bear expert to specify variables, relationships, and inputs into their models. The expert then adjusted the models until the outputs conformed to his expectations. In effect, the forecasts were the opinions of a single expert unaided by forecasting principles. Based on research to date, approaches based on unaided expert opinion are inappropriate to forecasting in situations with high complexity and much uncertainty.

Our audit of the second most relevant paper, H6, found that it was also based on faulty forecasting methodology. For example, it extrapolated nearly 100 years into the future on the basis of only five years of data – and data for these years were of doubtful validity.

In summary, experts' predictions, unaided by evidence-based forecasting procedures, should play no role in this decision. Without scientific forecasts of a substantial decline of the polar bear population *and* of net benefits from feasible policies arising from listing polar bears, a decision to list polar bears as threatened or endangered would be irresponsible.

Key words: adaptation, bias, climate change, decision making, endangered species, expert opinion, evaluation, evidence-based principles, expert judgment, extinction, forecasting methods, global warming, habitat loss, mathematical models, scientific method, sea ice.

Introduction

Polar bears have been described by some as the "canaries of climate change," and concern has been expressed over the survival of some sub-populations. We assessed the validity of long-term forecasts of selected polar bear populations by asking "Are the forecasts derived from accepted scientific procedures?"

We searched the Internet to identify scholarly publications on polar bears in order to assess whether population forecasts were consistent with proper forecasting principles. Second, we examined the references in the nine unpublished government reports written to support listing polar bears under the Endangered Species Act. Third, we examined the forecasting methods employed in two of those nine reports by assessing the procedures described in the reports against forecasting principles. We use the term "forecasting principles" to refer to guidelines on the selection of forecasting methods. The principles are based on evidence from scientific research that has revealed which methods provide the most accurate forecasts for a given situation.

Scientific forecasting procedures

Scientific research on forecasting has been conducted since the 1930s; important findings from the extensive literature on forecasting were first summarized in Armstrong (1978, 1985).

In the mid-1990s, the Forecasting Principles Project was established with the objective of summarizing all useful knowledge about forecasting. The evidence was codified as principles, or condition-action statements, to provide guidance on which methods to use in differing circumstances. The project led to the *Principles of Forecasting* handbook (Armstrong 2001). These principles were formulated by 40 internationally-recognized experts on forecasting methods and were reviewed by a further 123 leading experts on forecasting methods. The summarizing process alone was a four-year effort. We refer to the evidence-based methods as scientific forecasting procedures.

The strongest form of evidence is that derived from empirical studies that compare the performance of alternative methods. Ideally, "performance" is assessed by the ability of the selected method to provide useful *ex ante* forecasts. The weakest form of evidence is based on received wisdom about proper procedures. However, some of these principles seem self-evident (e.g., "Provide complete, simple and clear explanations of methods") and, as long as they were unchallenged by the available evidence, they were included. Some important principles are counter-intuitive: as a consequence, forecasts derived in ignorance of forecasting principles have no scientific standing.

The forecasting principles are available on forecastingprinciples.com, a site sponsored by the International Institute of Forecasters. The site claims to provide "all useful knowledge about forecasting" and asks visitors to submit any missing evidence. The site has been at the top of the list of sites in Internet searches for "forecasting" for many years. We expect about 200,000 unique visitors and 450,000 visits in 2007.

A summary of the principles, currently numbering 140, is provided as a checklist in the Forecasting Audit software available on the site. The strength of evidence is summarized briefly for each principle, and details are provided in Armstrong (2001) as well as in papers posted on the site.

General Assessment of Long-Term Polar Bear Population Forecasts

We conducted a Google Scholar Advanced Search on September 29, 2007, using the terms "polar bear" and either "forecast" or "predict." The search produced 997 unique sites out of 1,300 in total. We conducted a second search adding the term "forecasting principles" to the other terms, and found no sites. Similar results occurred when we added "forecastingprinciples.com," and in additional searches including the relevant publications on principles: "Principles of Forecasting"; "Long-range Forecasting"; "Forecasting for Environmental Decision-Making"; and "Makridakis." The last of these is the name of the first author of the largest selling text in forecasting and the book that has long been the top-selling book when one searches for "forecasting" on Amazon.com.

We also examined the references cited in the nine unpublished USGS Administrative Reports posted on the Internet at usgs.gov/newsroom/special/polar_bears/: Amstrup et al. (2007); Bergen et al (2007); DeWeaver (2007); Durner et al. (2007); Hunter et al. (2007); Obbard et al. (2007); Regehr et al. (2007); Rode et al. (2007) and Stirling et al. (2007). There were 654 references in total, with many of these cited in a number of the USGS Administrative Reports. We were unable to find any references that related to the validation of forecasting methods.

In short, we have been unable to find any support for the contention that polar bear forecasting efforts to date have followed accepted scientific principles.

We did find scientific papers that could be used as inputs to scientific forecasting procedures, of course. We cite a number of such papers in our audit. To ensure that we cited them properly, in December 2007, we sent a copy of our paper to all authors that we cited in a substantive manner, asking them to inform us if we have not properly referred to their findings. The authors did not object to the ways that we summarized their research.

Forecasting Audit of Two Key Papers Prepared to Support an Endangered Listing

We audited the forecasting procedures used in what we judged to be the two most crucial of the nine papers commissioned by the U. S. Department of the Interior to support the petition to classify polar bears as an endangered species.

The evidence-based principles upon which our audit was based were derived from many areas, including management, psychology, economics, politics, and weather, with the intention that they would apply to any type of forecasting problem. Some reviewers of our research have suggested that the principles do not apply to the physical sciences. We have asked for evidence to support that viewpoint, but have been unable to obtain responses. Readers can examine the principles and form their own judgments on this issue. For example, might one argue that the principle, "Ensure that information is reliable and that measurement error is low," does not apply because we are forecasting climate change?

In conducting the audits, each of three authors read the paper and independently rated the forecasting procedures described in it by using the Forecasting Audit software at forecastingprinciples.com. The rating scale ran from –2 to +2, the former referring to a violation of a principle and the latter signifying its proper application. After the initial round of ratings, we identified differences in our ratings and why they existed in an attempt to reach consensus. To the extent that we had difficulty in reaching consensus, we moved ratings toward "0".

Clearly forecasting audit ratings involve some subjectivity. Despite this, for each of the papers our ratings after the first round were in substantial agreement. Furthermore, we had little difficulty in reaching consensus by the third round.

In some cases, the papers did not provide sufficient details to allow for ratings. To resolve this issue, we contacted the authors of the two papers and requested further information. In addition, we asked them to review our ratings and to tell us whether they disagreed with any of them. In their reply, they refused to provide any responses to our requests. (See Note 2 at the end of our paper.)

Audit of AMD

We audited Amstrup, Marcot, and Douglas (2007) (henceforth AMD). That paper made forecasts of polar bear populations for 45, 75, and 100 years from the year 2000.

AMD implicitly assumed a complex chain of events. The causal chain was that: (1) global warming will occur; (2) this will both reduce the extent of and thin the summer sea ice; (3) polar bears will obtain less food by hunting from the sea ice platform than they do now; (4) they will not obtain adequate supplementary food using other means or from other sources; (5) the bear population will decline; (6) the designation of polar bears as an endangered species will solve the problem and will not have serious detrimental effects; and (7) there are no other policies that would produce better outcomes than those based on an endangered species classification.

AMD assumed that the general circulation models (GCMs) provide scientifically valid forecasts of global temperature and the extent and thickness of sea ice. They stated (AMD 2007, p. 2 and Fig 2 p. 83): "Our future forecasts are based largely on information derived from general circulation model (GCM) projections of the extent and spatiotemporal distribution of sea ice." That is, their forecasts are conditional on long-term global warming forecasts leading to a dramatic reduction in Arctic sea ice during maximum melt-back periods in spring, late summer and fall.

Green and Armstrong (2007) examined long-term climate forecasting efforts and were unable to find a single forecast of global warming that was based on scientific methods. The climate modelers' procedures violated many forecasting principles and some of the violations were critical. This formal auditing result is consistent with earlier cautions. For example, Soon et al. (2001) found that the current generation of GCMs is unable to meaningfully calculate the climatic effects of added atmospheric carbon dioxide given severe limitations from both the uncertainties and unknowns in representing all relevant physical processes. There appears to be much uncertainty about the direction and extent of global mean temperature changes in the long-term (Soon et al. 2001), the most plausible forecast is that mean temperatures will remain much the same.

The fact that the AMD forecasts rest on the GCM forecasts and that these forecasts lack a scientific basis (indeed, some climate modelers state that the GCM's do not provide forecasts), breaks the causal chain. Furthermore, the GCM models are not designed for analysis at a regional level as being considered by AMD and H6,

We audited AMD's polar bear population forecasting procedures to assess whether they would produce valid forecasts assuming valid climate and sea ice forecasts were available as inputs. Of the 140 forecasting principles, we agreed that 24 were irrelevant to the forecasting problem. We then examined principles on which our ratings differed. After two rounds of consultation (i.e., the process required three round in all), we were able to reach consensus on all 116 relevant principles. We found that AMD's procedures clearly violated 40 principles (Table 1) and appeared to violate 33 principles (Table 2). We were unable to rate 26 relevant principles (Table 3) due to a lack of information.

Table 1: Principles that were clearly violated in AMD

Setting Objectives:

1.2 Prior to forecasting, agree on actions to take assuming different possible forecasts.
1.3 Make sure forecasts are independent of politics.
1.4 Consider whether the events or series can be forecasted.
1.5 Obtain decision makers' agreement on methods.

Identify Data Sources:

3.5 Obtain information from similar (analogous) series or cases. Such information may help to estimate trends.

Collecting Data:

4.2 Ensure that information is reliable and that measurement error is low.

Selecting Methods:

6.1 List all the important selection criteria before evaluating methods.
6.2 Ask unbiased experts to rate potential methods.
6.7 Match the forecasting method(s) to the situation
6.8 Compare track records of various forecasting methods.
6.10 Examine the value of alternative forecasting methods.

Implementing Methods: General

7.3 Be conservative in situations of high uncertainty or instability.

Implementing Judgmental Methods:

8.1 Pretest the questions you intend to use to elicit judgmental forecasts.
8.2 Frame questions in alternative ways.
8.5 Obtain forecasts from heterogeneous experts.
8.7 Obtain forecasts from enough respondents.
8.8 Obtain multiple forecasts of an event from each expert.

Implementing Quantitative Methods:

9.1 Tailor the forecasting model to the horizon.
9.3 Do not use "fit" to develop the model.
9.5 Update models frequently.

Implementing Methods: Quantitative Models with Explanatory Variables:

10.6 Prepare forecasts for at least two alternative environments.
10.8 Apply the same principles to forecasts of explanatory variables.
10.9 Shrink the forecasts of change if there is high uncertainty for predictions of the explanatory variables.

Combining Forecasts:

12.1 Combine forecasts from approaches that differ.
12.2 Use many approaches (or forecasters), preferably at least five.
12.3 Use formal procedures to combine forecasts.
12.4 Start with equal weights.

Evaluating Methods:

13.6 Describe potential biases of forecasters.
13.10 Test assumptions for validity.
13.32 Conduct explicit cost-benefit analyses.

Assessing Uncertainty:

14.1 Estimate prediction intervals (PIs).
14.2 Use objective procedures to estimate explicit prediction intervals.
14.3 Develop prediction intervals by using empirical estimates based on realistic representations of forecasting situations.
14.5 Ensure consistency over the forecast horizon.
14.7 When assessing PIs, list possible outcomes and assess their likelihoods.
14.8 Obtain good feedback about forecast accuracy and the reasons why errors occurred.
14.9 Combine prediction intervals from alternative forecasting methods.
14.10 Use safety factors to adjust for overconfidence in the PIs.
14.11 Conduct experiments to evaluate forecasts.
14.13 Incorporate the uncertainty associated with the prediction of the explanatory variables in the prediction intervals.
14.14 Ask for a judgmental likelihood that a forecast will fall within a pre-defined minimum-maximum interval

Table 2: Principles that were apparently violated in AMD

Structuring the problem:

2.1 Identify possible outcomes prior to making forecasts.
2.7 Decompose time series by level and trend.

Identify Data Sources:

3.2 Ensure that the data match the forecasting situation.
3.3 Avoid biased data sources.
3.4 Use diverse sources of data.

Collecting Data:

4.1 Use unbiased and systematic procedures to collect data.
4.3 Ensure that the information is valid.

Selecting Methods:

6.4 Use quantitative methods rather than qualitative methods.
6.9 Assess acceptability and understandability of methods to users.

Implementing Methods: General

7.1 Keep forecasting methods simple.

Implementing Quantitative methods:

9.2 Match the model to the underlying phenomena.
9.4 Weight the most relevant data more heavily.

Implementing Methods: Quantitative Models with Explanatory Variables:

10.1 Rely on theory and domain expertise to select causal (or explanatory) variables.
10.2 Use all important variables.
10.5 Use different types of data to measure a relationship.

.

Combining Forecasts:

12.5 Use trimmed means, medians, or modes
12.7 Use domain knowledge to vary weights on component forecasts.
12.8 Combine forecasts when there is uncertainty about which method is best.
12.9 Combine forecasts when you are uncertain about the situation.
12.10 Combine forecasts when it is important to avoid large errors.

Evaluating Methods:

13.1 Compare reasonable methods.
13.2 Use objective tests of assumptions.
13.7 Assess the reliability and validity of the data.
13.8 Provide easy access to the data.
13.17 Examine all important criteria.
13.18 Specify criteria for evaluating methods prior to analyzing data.
13.27 Use ex post error measures to evaluate the effects of policy variables.

Assessing Uncertainty:

14.6 Describe reasons why the forecasts might be wrong.

Presenting Forecasts:

15.1 Present forecasts and supporting data in a simple and understandable form.
15.4 Present prediction intervals.

Learning That Will Improve Forecasting Procedures:

16.2 Seek feedback about forecasts.
16.3 Establish a formal review process for forecasting methods.

Table 3: Principles for which ratings could not be derived due to lack of information in AMD

Structuring the problem:

2.5 Structure problems to deal with important interactions among causal variables.

Collecting data:

4.4 Obtain all of the important data
4.5 Avoid the collection of irrelevant data

Preparing Data:

5.1 Clean the data.
5.2 Use transformations as required by expectations.
5.3 Adjust intermittent series.
5.4 Adjust for unsystematic past events.
5.5 Adjust for systematic events.
5.6 Use multiplicative seasonal factors for trended series when you can obtain good estimates for seasonal factors.
5.7 Damp seasonal factors for uncertainty

Selecting Methods:

6.6 Select simple methods unless empirical evidence calls for a more complex approach.

Implementing Methods: General

7.2 The forecasting method should provide a realistic representation of the situation

Implementing Judgmental Methods:

8.4 Provide numerical scales with several categories for experts' answers.

Implementing Methods: Quantitative Models with Explanatory Variables:

10.3 Rely on theory and domain expertise when specifying directions of relationships.
10.4 Use theory and domain expertise to estimate or limit the magnitude of relationships.

Integrating Judgmental and Quantitative Methods:

11.1 Use structured procedures to integrate judgmental and quantitative methods.
11.2 Use structured judgment as inputs to quantitative models.
11.3 Use pre-specified domain knowledge in selecting, weighting, and modifying quantitative methods.
11.4 Limit subjective adjustments of quantitative forecasts.

Evaluating Methods:

13.4 Describe conditions associated with the forecasting problem.
13.5 Tailor the analysis to the decision.
13.9 Provide full disclosure of methods.
13.11 Test the client's understanding of the methods.
13.19 Assess face validity.

Assessing Uncertainty:

14.12 Do not assess uncertainty in a traditional (unstructured) group meeting.

Learning That Will Improve Forecasting Procedures:

16.4 Establish a formal review process to ensure that forecasts are used properly.

In order to produce valid forecasts, it is necessary to use procedures that are consistent with accepted forecasting principles. AMD violated many forecasting principles, and so their forecasts would not be valid even if valid climate and ice forecasts were available. We describe some of the salient violations of forecasting principles below. Evidence for these and all principles is available in Chapter 20 of Armstrong (2001) and in the Forecasting Audit software at forecastingprinciples.com.

Match the forecasting method(s) to the situation (Principle 6.7)

The forecasts in AMD rely on the opinions of an expert who is knowledgeable in the domain. The opinions were transformed into a complex set of formulae, but were unaided by evidence-based forecasting principles.

Some studies (e.g., Tetlock 2005) suggest that judgmental forecasts by researchers who ignore accepted forecasting principles have little value in complex and uncertain situations. This apparently applies whether the opinions are expressed in words, spreadsheets, or mathematical models. It also applies regardless of how much scientific information is used by the experts. Among the reasons for this are:

a) Complexity: Individuals cannot assess complex relationships through unaided observations.
b) Coincidence: Individuals confuse correlation with causation.
c) Feedback: Individuals making judgmental predictions typically do not receive unambiguous feedback they can use to improve their forecasting.
d) Bias: Individuals have difficulty in obtaining or using evidence that contradicts their initial beliefs. This problem is especially serious among Individuals who view themselves as experts.

Despite the lack of validity of unaided forecasts by experts, many public policy decisions are based on their predictions. Research on persuasion has shown that people have substantial faith in the value of such forecasts and that faith increases when experts agree with one another. Although they may seem convincing at the time, expert forecasts can, a few years later, serve as important cautionary tales. Cerf and Navasky's (1998) book contains 310 pages of examples, such as Fermi Award-winning scientist John von Neumann's 1956 prediction that "A few decades hence, energy may be free". Examples of expert climate forecasts that turned out to be wrong are easy to find, such as UC Davis ecologist Kenneth Watt's prediction in a speech at Swarthmore College on Earth Day, April 22, 1970 that "If present trends continue, the world will be about four degrees colder in 1990, but eleven degrees colder in the year 2000. This is about twice what it would take to put us into an ice age."

Are such examples merely a matter of selective perception? The first author's review of empirical research on this problem led him to develop the "Seer-sucker Theory," which can be stated as "No matter how much evidence exists that seers do not exist, seers will find suckers" (Armstrong 1980). The amount of expertise does not matter beyond a basic minimum level. There are exceptions to the Seer-sucker Theory: experts can improve their forecasting when they receive well-summarized feedback on the accuracy of their forecasts and reasons why their forecasts were or were not accurate. This situation applies for short-term (up to five day) weather forecasts, but we are not aware of any such regime for long-term global climate forecasting. Even if there were such a regime, the feedback would trickle in over many years before it became useful for improving forecasting. Moreover, experts typically resist negative feedback and prefer to provide excuses for inaccurate forecasts (Tetlock 2005).

Research since 1980 has added support to the Seer-sucker Theory. In particular, Tetlock (2005) recruited 284 people whose professions included, "commenting or offering advice on political and economic trends." He asked them to forecast the probability that various situations would or would not occur, picking areas (geographic and substantive) within and outside their areas of expertise. By 2003, he had accumulated over 82,000 forecasts. The experts barely if at all outperformed non-experts and neither group did well against simple rules.

Comparative empirical studies have routinely concluded that judgmental forecasting by experts is the least accurate of the methods available to make forecasts. For example, Ascher

(1978, p. 200), in his analysis of long-term forecasts of electricity consumption, found that that was the case.

AMD also implicitly forecast—that is, they used their judgment unaided by scientific forecasting procedures—that a policy to classify polar bears as a threatened species would save the bears from future possible extinction. AMD did not include forecasts of the costs, planned and unintended, of such a policy.

Be conservative in situations of high uncertainty or instability (Principle 7.3)

Forecasts should be conservative when a situation is unstable, complex or uncertain. Being conservative means moving forecasts towards "no change" or, in cases that exhibit a well established long-term trend and where there is no reason to expect the trend to change, being conservative means moving forecasts toward the trend line. A long-term trend is one that has been evident over a period that is *much longer* than the period being forecast. Conservatism is a fundamental principle in forecasting.

The interaction between polar bears and their environment in the Arctic is complex and there is much uncertainty. For example, AMD associated warm temperatures with lower polar bear survival rates, yet cold temperatures have also been associated with similar outcomes, as this quote illustrates: "Abnormally heavy ice covered much of the eastern Beaufort Sea during the winter of 1973-1974. This resulted in major declines in numbers and productivity of polar bears and ringed seals in 1975" (Amstrup et al. 1986, p. 249).). Stirling (2002, p. 68 and 72) further expanded on the complexity of polar bear-sea-ice interactions:

> In the eastern Beaufort Sea, in years during and following heavy ice conditions in spring, we found a marked reduction in production of ringed seal pups and consequently in the natality of polar bears ... The effect appeared to last for about three years, after which productivity of both seals and bears increased again. These clear and major reductions in productivity of ringed seals in relation to ice conditions occurred at decadal-scale intervals in the mid-1970s and 1980s ... and, on the basis of less complete data, probably in the mid-1960s as well ... Recent analyses of ice anomalies in the Beaufort Sea have now also confirmed the existence of an approximately 10-year cycle in the region ... that is roughly in phase with a similar decadal-scale oscillation in the runoff from the Mackenzie River ... However, or whether, these regional-scale changes in ecological conditions have affected the reproduction and survival of young ringed seals and polar bears through the 1990s is not clear.

The inherent variability occurring at a regional scale adds to uncertainty. For example, Antarctic ice extent has been growing at the same time that sea and air temperatures have been increasing (e.g. Zhang 2007) while depth averaged oceanic temperatures around the Southeastern Bering Sea (Richter-Menge et al. 2007) have been undergoing relative cooling in 2006. Despite the warming of local air temperature by 1.6±0.6ºC, there was no sharp decline in the area over the continental shelf of the Canadian Beaufort Sea that was ice-covered for the 36 years from 1968 to 2003 (Melling et al. 2005).

Despite the uncertainty, instability, and complexity of the situation, AMD made predictions based on assumptions that we view as questionable. They also used little historical data.

Obtain forecasts from heterogeneous experts (Principle 8.5)

AMD's polar bear population forecasts were the product of a single expert. Experts vary in their knowledge and the way they approach problems, and bringing more information and different approaches to bear on a forecasting problem improves accuracy. When sufficient information is not available, forecasting can not be assumed valid. Also, in situations where experts might be biased, it is important to obtain forecasts from experts with different biases. Failing to follow this principle increases the risk that the forecasts obtained will be extreme when, in this situation, forecasts should be conservative (see Principle 7.3, above).

Use all important variables (Principle 10.2)

Dyck et al. (2007) recently noted that scenarios of polar bear decline grossly oversimplify the complex ecological relationships of the situation. In particular, AMD did not adequately consider the adaptability of polar bears. They mentioned the fact that polar bears evolved from brown bears 250,000 years ago (p. 2) but they appear to have ignored the fact that polar bears probably experienced much warmer conditions in the Arctic over that extended time period, with periods when sea ice habitat was less than those expected over the next century according to the GCM projections AMD have used. Several studies (Hamilton and Brigham-Grette 1991; Brigham-Grette and Hopkins 1995; Norgaard-Pedersen et al. 2007) have documented the dramatic reduction of sea ice in both the Northwest Alaskan coast and Northwest Greenland part of the Arctic Ocean during the very warm Interglacial of marine isotope stage 5e *ca.* 130,000 to 120,000 years ago. Brigham-Grette and Hopkins (1995, p. 159) noted that the "winter sea-ice limit was north of Bering Strait, at least 800 km north of its present position, and the Bering Sea was perennially ice-free" and that "[the more saline] Atlantic water may have been present on the shallow Beaufort Shelf, suggesting that the Arctic Ocean was not stratified and the Arctic sea-ice cover was not perennial for some period." On the face of it, the nature and extent of polar bear adaptability seems crucial to any forecasts that assume dramatic changes in the bears' environment.

AMD's forecasts were commissioned to inform public policy decisions, but they do not explicitly forecast the effects of different policies. For example, in the event of the polar bear population coming under stress due to inadequate summer food, what would be the costs and effects of creating conservation [protected?] areas where marine and land-based activities were prohibited at critical seasons? In addition, what would be the costs and benefits of a smaller but stable population of polar bears in some polar sub-regions? And how would the net costs of such alternative policies compare with the net costs of listing polar bears?

Make sure forecasts are independent of politics (Principle 1.3)

By politics, we refer to any type of organizational biases or pressures. While different stakeholders may prefer particular forecasts, if forecasters are influenced by such considerations, forecast accuracy will suffer. The Executive Summary document[1] noted that "the Secretary of the Interior asked the U.S. Geological Survey (USGS) to generate new scientific data, models, and interpretations on polar bears and their sea ice habitats, to support the "U.S. Fish and Wildlife service polar bear listing decision" (http://www.doi.gov/news/06_News_Releases/061227.html). The authors of the AMD administrative report are all employees of the U.S. government agencies that are trying to support this decision.

Audit of Hunter et al (H6)

Hunter et al. (2007), which we refer to here as H6, forecasted polar bear numbers in the southern Beaufort Sea for 45, 75, and 100 years from 2000. To do so, they implicitly assumed the following causal chain: (1) global warming will occur; (2) frequent "bad years" will be a consequence of global warming; (3) polar bears will not adapt to "bad years"; (4) the population of polar bears will decline dramatically from negative effects of "bad years" alone; (5) the

[1] http://www.usgs.gov/newsroom/special/polar_bears/docs/executive_summary.pdf

designation of polar bears as an endangered species will solve the problem and will not have serious detrimental effects; and (6) there are no other policies that would produce better outcomes than those based on listing polar bears under the Endangered Species Act.

Like AMD, H6 also accepted GCM forecasts of global warming and reduced extent and thickness of sea ice. They stated that "we extracted forecasts of the availability of sea ice for polar bears in the SB [southern Beaufort Sea] region, using monthly forecasts of sea ice concentrations from 10 IPCC Fourth Assessment Report (AR4) fully-coupled general circulation models" (p. 11 of H6). That is, their forecasts are conditional on long-term forecasts of global warming producing dramatic effects. However, Green and Armstrong (2007) were unable to find any scientifically-valid forecasts to support the hypothesized predictions of global warming throughout the 21st Century.

We nevertheless audited H6's polar bear population forecasting procedures to assess whether they would produce valid forecasts if valid climate and sea ice forecasts were available as inputs.

Each of the authors read H6 and independently rated the forecasting procedures described in it using the Forecasting Audit software at forecastingprinciples.com. Of the 140 forecasting principles, we agreed that 35 were irrelevant to the forecasting problem. We then examined principles on which our ratings differed, and after three rounds of consultation we were able to reach consensus on all 105 relevant principles. To the extent that we had difficulty in reaching consensus, we moved ratings toward "0".-

We found that H6's procedures clearly violated 61 principles (Appendix Table A) and appeared to violate an additional 19 principles (Appendix Table B). We were unable to rate 15 relevant principles (Appendix Table C) due to a lack of information.

Given that many of the violations in H6 were similar to those in AMD, we provide the H6 audit details in the appendix. Here are some examples of clear violations of principles, some of which are, on their own, sufficient to render the forecasts useless:

Decisions, actions, and biases (Principles 1.1 – 1.3)

The H6 authors improperly formulated the problem for forecasting. They did not describe alternative decisions that might be taken (1.1), nor did they propose relationships between possible forecasts and alternative decisions (1.2). For example, what decision would be implied by a forecast that bear numbers will increase to the point where they become a menace to existing human settlements? These problems relate to the biased manner in which the problem was stated: "USGS science strategy to support U.S. Fish and Wildlife Service polar bear listing decision" (1.3). Research is often prone to bias, sometimes due to unknown preferences or interests, but it is nevertheless important to try to avoid it, and it is clearly improper to undertake a research project on the understanding that there is a desired finding.

Ensure that information is reliable and that measurement error is low (Principle 4.2)

Long-term forecasts require enormous amounts of valid and reliable data. Armstrong (1985, p. 166) refers to two rules of thumb for how much data are needed for extrapolating *h* years ahead. One calls for $4h^{1/2}$ years of historical data and the other calls for *h* years. These rules imply that H6 should have based any extrapolations on 40 to 100 years of historical data. In order to forecast using causal methods, it is necessary to have reliable data over sufficiently long periods for all variables to have varied relative to each other on a large number of occasions.

H6 clearly violates this principle with its reliance on five years of data with unknown measurement errors. We were not convinced that the capture data on which they rely could provide representative samples of bears in the southern Beaufort Sea given the vast area involved and difficulties in spotting and capturing the bears. Moreover, bears wander over long distances and do not respect administrative boundaries (Amstrup et al. 2004). The validity of the data is likely to be compromised further by imposing a speculative demographic model on the raw capture-recapture data (Amstrup et al. 2001; Regehr et al. 2006).

Be conservative in situations of high uncertainty or instability (Principle 7.3)

The report violates the principle of being conservative where there is uncertainty.

The situation regarding polar bears in the southern Beaufort Sea is complex and there is much uncertainty. For example, on the basis of five years of data, H6 associated warm temperatures with lower polar bear survival rates, yet as noted earlier, cold temperatures have also been associated with similar outcomes: "Abnormally heavy ice covered much of the eastern Beaufort Sea during the winter of 1973-1974. This resulted in major declines in numbers and productivity of polar bears and ringed seals in 1975" (Amstrup et al. 1986, p. 249).

We repeat what we stated earlier, that regional changes add to uncertainty, noting the Antarctic ice extent has been growing at the same time that sea and air temperatures have been increasing (e.g. Zhang 2007) while depth averaged oceanic temperatures around the southeastern Bering Sea have been undergoing relative cooling in 2006 (Richter-Menge et al. 2007). Despite the warming of local air temperature by 1.6±0.6ºC, there was no sharp decline in the area over the continental shelf of the Canadian Beaufort Sea that was covered in ice for the 36 years from 1968 to 2003 (Melling et al. 2005).

Tailor the forecasting model to the horizon (Principle 9.1)

When forecasting over the long term, as in H6, forecasting models should be based on long-term trends. In contrast, the H6 authors built models based entirely on estimates derived from only five years of recent data.

Update frequently (Principle 9.5)

H6 did not include the most recent year, 2006, when estimating their model. From the supplementary information provided in Figure 3 of Regehr et al. (2007), one finds that the number of ice-free days for the 2006 season was about 105: close to the mean of the "good" ice years.

The latest "Alaska Marine Mammal Stock Assessment, 2006" report by Angliss and Outlaw (2007, p. 218), states that

> The Southern Beaufort Sea [polar bear] Stock is not classified as "depleted' under the MMPA or listed as 'threatened' or 'endangered' under terms of the Endangered Species Act. This stock is assumed to be within optimum sustainable population levels.

Use all important variables (Principle 10.2)

With causal models, it is important to incorporate policy variables if they might vary or if the purpose is to decide what policy to implement. H6 did not include policy variables such as seasonal protection of bears' critical habitat, or changes to hunting rules.

Other variables should also be included, such as migration, snow conditions, and windiness. For example Holloway and Sou (2002), Ogi and Wallace (2007), and Nghiem et al. (2007) suggested that large-scale atmospheric winds and related patterns play an important role in causing both the decline in extent and thinning of Arctic sea ice; those effects were not included in the GCM forecasts of sea ice (and hence the quality of polar bear habitat).

In addition, Dyck et al. (2007) recently noted that future scenarios of polar bear decline grossly oversimplify the complex ecological relationships of the situation. This is why the extent and kind of polar bear adaptability is crucial to any forecasts that assume dramatic changes in the bears' environment.

Use different types of data to measure a relationship (Principle 10.5)

This principle is important when there is uncertainty about the magnitudes of relationships between causal variables (such as ice extent) and the event being forecast (polar bear population) and when large changes are expected in the causal variables. In the case of the latter condition, H6 accept the GCM model predictions of large declines in summer ice throughout the 21st century, so their forecasts are very sensitive to their estimate of the magnitude of the effect of ice extent on bears. Yet H6 base their estimate of this important relationship on only five years of data. They might, for example, have independently estimated the magnitude of the relationship by obtaining estimates of polar bear populations during much warmer and much colder periods in the past. The supplementary information from Figure 3 of Regehr et al. (2007) shows that 1987, 1993 and 1998 were exceptional seasons with the number of ice-free days longer than 150 days (i.e., significantly above the 135 ice-free days documented for 2004-2005) in the southern Beaufort sea, yet there were no apparent negative impacts on the polar bear population and wellbeing (see for example Amstrup et al. 2001).

Match the model to the underlying phenomena (Principle 9.2)

It is important for the readers to know what is meant by “Southern Beaufort Sea” (SB) in the H6 report because of the poor spatial resolution of the GCMs. H6 states: “Because GCMs do not provide suitable forecasts for areas as small as the SB, we used sea ice concentration for a larger area composed of 5 IUCN polar bear management units (Aars et al. 2006) with ice dynamics similar to the SB management unit (Barents Sea, Beaufort Sea, Chukchi Sea, Kara Sea and Laptev Sea; see Rigor and Wallace 2004, Durner et al. 2007). We assumed that the general trend in sea ice availability in these 5 units was representative of the general trend in the Southern Beaufort region.” (p. 12). Given the unique ecological, geographical, meteorological, and climatological conditions in each of the five circumpolar seas, we did not find this assumption plausible.

When assessing prediction intervals (PIs), list possible outcomes and assess their likelihoods (Principle 14.7)

To assess meaningful PIs, it helps to think of diverse possible outcomes. The H6 authors did not appear to consider, for example, the possibility that polar bears might adapt to terrestrial life over summer months by finding alternative food sources (such as is the case in the Southern Hudson Bay populations, or elsewhere; see references in Stempniewicz 2006; Dyck and Romberg 2007) or by successfully congregating in smaller or localized ice-hunting areas. Consideration of these and other possible adaptations and outcomes would have likely led the H6 authors to be less confident (provide wider prediction intervals) about a bad outcome for bears. Extending this exercise to the forecasts of climate and summer ice extent would have further widened the range of other outcomes.

Conclusions

The issue of listing a species under the Endangered Species Act should be based on credible scientific forecasts.

Based on our Internet search of the published scholarly research and on appeals to other researchers we have been unable to locate any papers that referred to scientific procedures for making forecasts of polar bear populations. Furthermore, a review of the references in the nine government reports written to support the listing of polar bears under the Endangered Species Act failed to find any papers relevant to scientific forecasting procedures.

We take no issue with the scientific work of the researchers whose work we have reviewed as it relates to the past. Our concern is that there are currently no scientific forecasts of the polar bear population; nor of direction or magnitude of changes. Without scientific forecasts of a substantial decline of the polar bear population *and* of net benefits from feasible policies arising from listing polar bears, a decision to list polar bears as threatened or endangered would be irresponsible.

Appendix

Table A: Principles that were clearly violated in H6

1. Setting Objectives:

1.3 Make sure forecasts are independent of politics.
1.4 Consider whether the events or series can be forecasted.

2. Structuring the problem:

2.6 Structure problems that involve causal chains.

3. Identify Data Sources:

3.4 Use diverse sources of data.
3.5 Obtain information from similar (analogous) series or cases. Such information may help to estimate trends.

4. Collecting Data:

4.4 Obtain all of the important data

5. Preparing Data:

5.2 Use transformations as required by expectations.
5.4 Adjust for unsystematic past events.
5.5 Adjust for systematic events.

6. Selecting Methods:

6.1 List all the important selection criteria before evaluating methods.
6.2 Ask unbiased experts to rate potential methods.
6.6 Select simple methods unless empirical evidence calls for a more complex approach.
6.7 Match the forecasting method(s) to the situation.
6.8 Compare track records of various forecasting methods.
6.10 Examine the value of alternative forecasting methods.

7. Implementing Methods: General

7.1 Keep forecasting methods simple.
7.2 The forecasting method should provide a realistic representation of the situation.
7.3 Be conservative in situations of high uncertainty or instability.
7.4 Do not forecast cycles.

9. Implementing Quantitative Methods:

9.1 Tailor the forecasting model to the horizon.
9.2 Match the model to the underlying phenomena.
9.3 Do not use "fit" to develop the model.
9.5 Update models frequently.

10. Implementing Methods: Quantitative Models with Explanatory Variables:

10.2 Use all important variables.
10.5 Use different types of data to measure a relationship.
10.7 Forecast for alternate interventions.
10.9 Shrink the forecasts of change if there is high uncertainty for predictions of the explanatory variables.

11. Integrating Judgmental and Quantitative Methods:

11.1 Use structured procedures to integrate judgmental and quantitative methods.
11.2 Use structured judgment as inputs to quantitative models.
11.3 Use pre-specified domain knowledge in selecting, weighting, and modifying quantitative methods.

12. Combining Forecasts:

12.1 Combine forecasts from approaches that differ.
12.2 Use many approaches (or forecasters), preferably at least five.
12.3 Use formal procedures to combine forecasts.
12.8 Combine forecasts when there is uncertainty about which method is best.
12.9 Combine forecasts when you are uncertain about the situation.
12.10 Combine forecasts when it is important to avoid large errors.

13. Evaluating Methods:

13.1 Compare reasonable methods.
13.2 Use objective tests of assumptions.
13.3 Design test situations to match the forecasting problem.

13.5 Tailor the analysis to the decision.
13.6 Describe potential biases of forecasters.
13.7 Assess the reliability and validity of the data.
13.8 Provide easy access to the data.
13.10 Test assumptions for validity.
13.12 Use direct replications of evaluations to identify mistakes.
13.13 Replicate forecast evaluations to assess their reliability.
13.16 Compare forecasts generated by different methods.
13.17 Examine all important criteria.
13.18 Specify criteria for evaluating methods prior to analyzing data.
13.26 Use out-of-sample (ex ante) error measures.
13.27 Use ex post error measures to evaluate the effects of policy variables.
13.31 Base comparisons of methods on large samples of forecasts.

14. Assessing Uncertainty:

14.3 Develop prediction intervals by using empirical estimates based on realistic representations of forecasting situations.
14.5 Ensure consistency over the forecast horizon.
14.9 Combine prediction intervals from alternative forecasting methods.
14.10 Use safety factors to adjust for overconfidence in the PIs.
14.11 Conduct experiments to evaluate forecasts.
14.13 Incorporate the uncertainty associated with the prediction of the explanatory variables in the prediction intervals.
14.14 Ask for a judgmental likelihood that a forecast will fall within a pre-defined minimum-maximum interval (not by asking people to set upper and lower confidence levels).

15. Presenting Forecasts:

15.1 Present forecasts and supporting data in a simple and understandable form.
15.2 Provide complete, simple, and clear explanations of methods.

Table B: Principles that were apparently violated in H6

1. Setting Objectives:

1.1 Describe decisions that might be affected by the forecasts.
1.2 Prior to forecasting, agree on actions to take assuming different possible forecasts.

2. Structuring the problem:

2.1 Identify possible outcomes prior to making forecasts.
2.3 Decompose the problem into parts.

3. Identify Data Sources:

3.2 Ensure that the data match the forecasting situation.
3.3 Avoid biased data sources.

4. Collecting Data:

4.2 Ensure that information is reliable and that measurement error is low.
4.3 Ensure that the information is valid.

5. Preparing Data:

5.3 Adjust intermittent series.
5.7 Damp seasonal factors for uncertainty
5.8 Use graphical displays for data.

7. Implementing Methods: General

7.6 Pool similar types of data.

10. Implementing Methods: Quantitative Models with Explanatory Variables:

10.4 Use theory and domain expertise to estimate or limit the magnitude of relationships.
10.8 Apply the same principles to forecasts of explanatory variables.

13. Evaluating Methods:

13.4 Describe conditions associated with the forecasting problem.
13.9 Provide full disclosure of methods.

14. Assessing Uncertainty:

14.6 Describe reasons why the forecasts might be wrong.
14.7 When assessing PIs, list possible outcomes and assess their likelihoods.
14.8 Obtain good feedback about forecast accuracy and the reasons why errors occurred.

Table C: Principles for which ratings could not be derived due to lack of information in H6

1. Setting Objectives:

1.5 Obtain decision makers' agreement on methods

2. Structuring the problem:

2.7 Decompose time series by level and trend

3. Identify Data Sources:

3.1 Use theory to guide the search for information on explanatory variables

4. Collecting Data:

4.1 Use unbiased and systematic procedures to collect data
4.5 Avoid the collection of irrelevant data

5. Preparing Data:

5.1 Clean the data

6. Selecting Methods:

6.4 Use quantitative methods rather than qualitative methods
6.5 Use causal methods rather than naive methods if feasible
6.9 Assess acceptability and understandability of methods to users

13. Evaluating Methods:

13.11 Test the client's understanding of the methods
13.19 Assess face validity

15. Presenting Forecasts:

15.3 Describe your assumptions

16. Learning That Will Improve Forecasting Procedures:

16.2 Seek feedback about forecasts
16.3 Establish a formal review process for forecasting methods
16.4 Establish a formal review process to ensure that forecasts are used properly

Notes:

1) Our interest in the topic of this paper was piqued when the State of Alaska hired us as consultants in late-September 2007 to assess forecasts that had been prepared "to Support U.S. Fish and Wildlife Service Polar Bear Listing Decision." We were impressed by the importance of the issue and, after providing our assessment, we decided to continue working on it and to prepare a paper for publication. These latter efforts have not been funded. We take responsibility for all judgments and for any errors that we might have made.

2) On November 27, 2007, we sent a draft of our paper to the authors of the U.S. Geological Service administrative papers that we audited and stated:

> As we note in our paper, there are elements of subjectivity in making the audit ratings. Should you feel that any of our ratings were incorrect, we would be grateful if you would you provide us with evidence that would lead to a different assessment. The same goes for any principle that you think does not

apply, or to any principles that we might have overlooked. There are some areas that we could not rate due to a lack of information. Should you have information on those topics, we would be interested. Finally, we would be interested in peer review that you or your colleagues could provide, and in suggestions on how to improve the accuracy and clarity of our paper.

We received a reply from Steven C. Amstrup on November 30, 2007 that said: "We all decline to offer preview comments on your attached manuscript. Please feel free, however, to list any of us as potential referees when you submit your manuscript for publication."

3) We invite others to conduct forecasting audits of AMD, H6, any of the other papers prepared to support the endangered species listing, or any other papers relevant to long-term forecasting of the polar bear population. Note that the audit process calls for two or more raters. The audits can be submitted for publication on pubicpolicyforecasting.com along with the auditors' bios and any information relevant potential sources of bias.

4) We seek information about scientifically developed forecasting studies, published or unpublished, that are relevant to polar bear forecasting.

5) We seek further peer review on this paper.

Acknowledgments

We thank Milton Freeman, Paul Goodwin, and Tom Stewart for their comments on earlier drafts. Janice Dow provided editorial assistance.

References

Amstrup, S.C, Marcot, B.G. and Douglas, D.C. (2007). Forecasting the rangewide status of polar bears at selected times in the 21st century. USGS Alaska Science Center, Anchorage, Administrative Report.

Amstrup, S.C., McDonald, T.L. and Durner, G.M. (2004). Using satellite radiotelemetry data to delineate and manage wildlife populations. Wildlife Society Bulletin, 32, 661-679.

Amstrup, S.C., McDonald, T.L. and Stirling, I. (2001). Polar bears in the Beaufort Sea: A 30-year mark-recapture case history. Journal of Agricultural, Biological, and Environmental Statistics, 6, 221-234.

Amstrup, S.C., Stirling, I. and Lentfer, J.W. (1986). Past and present status of polar bears in Alaska. *Wildlife Society Bulletin*, *14*, 241-254.

Angliss, R.P., and Outlaw, R.B. (2007). Alaska marine mammal stock assessments, 2006. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-168, 244 p. Available at http://www.nmfs.noaa.gov/pr/pdfs/sars/ak2006.pdf

Armstrong, J.S. (1978; 1985). *Long-Range Forecasting: From Crystal Ball to Computer.* New York: Wiley-Interscience.

Armstrong, J.S. (1980). The Seer-sucker theory: The value of experts in forecasting. *Technology Review 83* (June-July), 16-24.

Armstrong, J.S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Kluwer Academic Publishers.

Ascher W. 1978. *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Bergen, S., Durner, G. M., Douglas, D. C. and Amstrup, S. C. (2007). Predicting movements of female Polar bears between summer sea ice foraging habitats and terrestrial denning habitats of Alaska in the 21st Century: Proposed Methodology and Pilot Assessment, USGS Alaska Science Center, Anchorage, Administrative Report.

Brigham-Grette, J., and Hopkins, D.M. (1995). Emergent marine record and paleoclimate of the Last Interglaciation along the Northwest Alaskan coast. *Quaternary Research*, 43, 159-173.

Cerf, C. and Navasky, V. (1998). *The Experts Speak*. New York: Pantheon. Baltimore, MD: Johns Hopkins University Press.

DeWeaver, E. (2007). Uncertainty in climate model projections of Arctic sea ice decline: An evaluation relevant to polar bears. USGS Alaska Science Center, Anchorage, Administrative Report.

Durner, G.M., Douglas, D.C. Nielson, R.M. Amstrup, S.C. and McDonald, T.L. (2007). Predicting the future distribution of polar bear habitat in the Polar Basin from resource selection functions applied to 21st century general circulation model projections of sea ice. USGS Alaska Science Center, Anchorage, Administrative Report.

Dyck, M.G., and Romberg, S. (2007). Observations of a wild polar bear (*Ursus maritimus*) successfully fishing Arctic charr (*Salvelinus alpinus*) and Fourhorn sculpin (*Myoxocephalus quadricornis*). *Polar Biology*, 30, 1625-1628.

Dyck, M.G., Soon, W., Baydack, R.K., Legates, D.R., Baliunas, S., Ball, T.F., and Hancock, L.O. (2007). Polar bears of western Hudson Bay and climate change: Are warming spring air temperatures the "ultimate" survival control factor? *Ecological Complexity*, 4, 73-84.

Green, K.C. and Armstrong, J.S. (2007), Global warming: forecasts by scientists versus scientific forecasts. *Energy and Environment*, 18, 997-1021.

Hamilton, T.D., and Brigham-Grette, J. (1991). The last interglaciation in Alaska: Stratigraphy and paleoecology of potential sites. *Quaternary International*, 10-12, 49-71.

Holloway, G., Sou, T. (2002). Has Arctic sea ice rapidly thinned? *Journal of Climate*, 15, 1691-1701.

Hunter, C.M., Caswell, H., Runge, M.C., Amstrup, S.C., Regehr, E.V. and Stirling I. (2007). Polar bears in the Southern Beaufort Sea II: Demography and population growth in relation to sea ice conditions. USGS Alaska Science Center, Anchorage, Administrative Report.

Makridakis, S., Wheelwright, S.C., and Hyndman, R.J. (1998). *Forecasting: Methods and Applications* (3rd ed.), Hoboken, NJ: John Wiley.

Melling, H., Riedel, D.A., and Gedalof, Z. (2005). Trends in the draft and extent of seasonal pack ice, Canadian Beaufort Sea. *Geophysical Research Letters*, 32, L24501, doi:10.1029/2005GL024483.

Nghiem, S.V., Rigor, I.G., Perovich, D.K., Clemente-Colon, P., Weatherly, J.W., and Neumann, G. (2007). Ra[id reduction of Arctic perennial sea ice. *Geophysical Research Letters*, 34, L19504, doi:10.1029/2007GL031138

Norgaard-Pedersen, N., Mikkelsen, N., Lassen, S.J., Kristoffersen, Y., and Sheldon, E. (2007). Reduced sea ice concentrations in the Arctic Ocean during the last interglacial period revealed by sediment cores off northern Greenland. *Paleoceanography*, 22, PA1218, doi:10.1029/2006PA001283.

Obbard, M.E., McDonald, T.L. Howe, E.J. Regehr, E.V. and Richardson, E.S. (2007). Trends in abundance and survival for polar bears from Southern Hudson Bay, Canada, 1984-2005. USGS Alaska Science Center, Anchorage, Administrative Report.

Ogi, M., and Wallace, J.M. (2007). Summer minimum Arctic sea ice extent and the associated summer atmospheric circulation. *Geophysical Research Letters*, 34, L12705, doi:10.1029/2007GL029897

Regehr, E.V., Hunter, C.M. Caswell, H. Amstrup, S.C. and Stirling, I. (2007). Polar bears in the Southern Beaufort Sea I: Survival and breeding in relation to sea ice conditions, 2001-2006. USGS Alaska Science Center, Anchorage, Administrative Report.

Richter-Menge, J. and Coauthors (2007). State of the climate in 2006: Arctic. *Bulletin of the American Meteorological Society*, 88, S62-S71.

Rode, K. D., Amstrup, S. C., and Regehr, E.V. (2007), Polar bears in the Southern Beaufort Sea III: Stature, mass, and cub recruitment in relationship to time and sea ice extent between 1982 and 2006. USGS Alaska Science Center, Anchorage, Administrative Report.

Soon, W., Baliunas, S., Idso, S. B., Kondratyev, K. Ya., Posmentier, E. S. (2001). Modeling climatic effects of anthropogenic carbon dioxide emissions: Unknowns and uncertainties. *Climate Research*, 18, 259 – 275.

Stempniewicz, L. 2006. Polar bear predatory behaviour toward moulting barnacle geese and nesting glaucous gulls on Spitsbergen. *Arctic*, 59. 247-251.

Stirling, I. (2002). Polar bears and seals in the Eastern Beaufort Sea and Amundsen Gulf: A synthesis of population trends and ecological relationships over three decades. *Arctic*, 55 (Suppl. 1), 59-76.

Stirling, I., McDonald, T.L. Richardson, E.S. and Regehr, E.V. (2007). Polar bear population status in the Northern Beaufort Sea. USGS Alaska Science Center, Anchorage, Administrative Report.

Tetlock, P.E. (2005). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

Zhang, J. (2007). Increasing Antarctic sea ice under warming atmospheric and oceanic conditions. *Journal of Climate*, 20, 2515-2529.