Polar Bear Population Forecasts: A Public-Policy Forecasting Audit

Working Paper Version 68: March 28, 2008

J. Scott Armstrong
The Wharton School, University of Pennsylvania
747 Huntsman, Philadelphia, PA 19104
armstrong@wharton.upenn.edu; jscottarmstrong.com; T +1 610 622 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
wsoon@cfa.harvard.edu; T +1 617 495 7488

Abstract

Calls to list polar bears as a threatened species under the U.S. Endangered Species Act are based on forecasts of substantial long-term declines in their population. Nine government reports were prepared to support the listing decision. We assessed these reports in light of evidence-based (scientific) forecasting principles. None referred to works on scientific forecasting methodology. Of the nine, Amstrup, Marcot and Douglas (2007) and Hunter et al. (2007) were the most relevant to the listing decision. Their forecasts were products of complex sets of assumptions. The first in both cases was the erroneous assumption that General Circulation Models provide valid forecasts of summer sea ice in the regions inhabited by polar bears. We nevertheless audited their conditional forecasts of what would happen to the polar bear population *assuming*, as the authors did, that the extent of summer sea ice would decrease substantially over the coming decades. We found that Amstrup et al. properly applied only 15% of relevant forecasting principles and Hunter et al. only 10%. We believe that their forecasts are unscientific and should therefore be of no consequence to decision makers. We recommend that *all* relevant principles be properly applied when important public policy decisions depend on accurate forecasts.

Key words: adaptation, bias, climate change, decision making, endangered species, expert opinion, extinction, 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." Despite widespread agreement that the polar bear population rose rapidly over recent years after the imposition of stricter hunting rules (Prestrud and Stirling 1994), new concerns have been expressed that climate change will threaten the survival of some sub-populations in the 21st Century. Such concerns led the U.S. Fish and Wildlife Service to consider listing polar bears as a threatened species under the Endangered Species Act. To list a species that is currently in good health must surely require valid forecasts that its population would, if it were not listed, decline to levels that threaten the viability of the species. The decision on listing polar bears thus rests on long-term forecasts.

In order to provide the necessary forecasts for the listing decision, the Secretary of the Interior and the Fish and Wildlife Service requested that the U.S. Geological Survey conduct additional analyses. The Geological Survey in turn commissioned nine "administrative reports" to satisfy the request.

We asked "Are the forecasts derived from accepted scientific procedures?" In order to answer this question, we first examined the references in the nine unpublished government reports. Second, we assessed the forecasting procedures described in two of those nine reports against forecasting principles. The forecasting principles are derived from evidence obtained from scientific research that has revealed which methods provide the most accurate forecasts for a given situation and which methods should be avoided.

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 under different circumstances. The project led to the *Principles of Forecasting* handbook (Armstrong 2001). The principles were formulated by forty internationally-recognized experts on forecasting methods and were reviewed by 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 which is 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, it is reasonable that decision makers and the public should expect people who make forecasts to be familiar with the principles of forecasting just as a patient expects his physician to be familiar with the procedures dictated by medical science.

The principles were derived from many fields, including demography, economics, engineering, finance, management, medicine, psychology, politics, and weather in order to ensure that all relevant evidence was taken into account and so that the principles would apply to all types of forecasting problem. Some people who have commented on our research suggested that the principles do not apply to the physical sciences. We have asked reviewers for evidence to support that assertion, but have been unable to obtain useful 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 when forecasting polar bear numbers?

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 Forecasting Principles site has been at the top of the list of sites in Internet searches for "forecasting" for many years.

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 examined all references cited in the nine unpublished U.S. Geological Survey Administrative Reports posted on the Internet at http://usgs.gov/newsroom/special/polar_bears/. The reports were 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). The reports included 444 unique references in total. We were unable to find any references to works providing evidence that the forecasting methods used in the reports had been previously validated.

Forecasting Audit of Key Reports Prepared to Support the Listing of Polar Bears

We audited the forecasting procedures described in the reports that we judged provided the strongest support, in the form of forecasts, for listing polar bears. These were Amstrup et al. (2007), which we refer to as AMD, and Hunter et al. (2007), which we refer to as H6. The reports each provided forecasts of polar bear populations for 45, 75, and 100 years from the year 2000.

The reports make recommendations with respect to the polar bear listing decision. These do not follow logically from their research in that they simply made forecasts of the polar bear population. To go from the forecasts to policy recommendations, the following assumptions are required:

(1) "global warming" will occur and will reduce the extent of summer sea ice;
(2) polar bears will not adapt and will thus obtain less food by hunting from the sea ice platform than they do now;
(3) listing polar bears as a threatened or endangered species will result in policies that will solve the problem without serious detrimental effects; and
(4) other policies would be inferior to those based on the Endangered Species Act.

Regarding the first of these issues, AMD and H6 assumed that general circulation models (GCMs) provide scientifically valid forecasts of global temperature and the extent and thickness of sea ice. They stated (AMD, p. 2 and Fig 2 on 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." H6 stated that "we extracted forecasts of the availability of sea ice for polar bears in the 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, the forecasts of both AMD and H6 are conditional on long-term global warming leading to a dramatic reduction in Arctic sea ice during 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. Their audit of the GCM climate modelers' procedures properly followed only 13% of the relevant forecasting principles and some of the contraventions of principles were critical. This finding was consistent with earlier cautions. For example, Soon et al. (2001) found that the current generation of GCMs is unable to meaningfully calculate the effects on climate of additional atmospheric carbon dioxide given the severe limitations imposed by uncertainty about the past and present climate and ignorance about relevant weather and climate processes. Indeed, some climate modelers state that the GCMs do *not* provide forecasts. Here is a statement by one of the lead authors of the IPCC's AR4:

> '…there are no predictions by IPCC at all. And there never have been. The IPCC instead proffers "what if" projections of future climate that correspond to certain emissions scenarios. There are a number of assumptions that go into these emissions scenarios. They are intended to cover a range of possible self consistent "story lines" that then provide decision makers with information about which paths might be more desirable.'
>
> (Written by Kevin Trenberth of the Climate Analysis Section, National Center for Atmospheric Research, and posted on ClimateFeedback at nature.com on June 4, 2007.)

No scientific evidence was provided to support AMD's and H6's assumptions about any of the four vital issues we identified above. Given the lack of good evidence on vital issues, the two administrative reports are of no value to decision makers.

We nevertheless audited AMD's and H6's polar bear population forecasting procedures to assess whether they would produce valid forecasts if the underlying assumptions had been valid.

In conducting the audits, the three of us read each report and independently rated the forecasting procedures described in it by using the Forecasting Audit software at forecastingprinciples.com. The rating scale runs from –2 to +2, with the former indicating the procedures contravene the principle and the latter signifying that it is properly applied. After the initial round of ratings, we examined differences in our ratings in an attempt to reach consensus. To the extent that we had difficulty in reaching consensus, we moved ratings toward "0." When we were not all agreed that a principle was clearly contravened, we gave a rating of -1 or "apparently contravened."

Here is an example of a principle that was contravened in both of the reports we audited, and also in the other seven reports:

> *Make sure forecasts are independent of politics (Principle 1.3)*
>
> By politics, we mean any type of organizational bias or pressure. It is not unusual for different stakeholders to prefer particular forecasts but, if forecasters are influenced by such considerations, forecast accuracy may suffer. The U.S. Geological Survey 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."[2]

---

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

[2] See http://www.doi.gov/news/06_News_Releases/061227.html

While it was easy to code the reports' procedures against the above principle, the ratings were subjective for many principles. Despite the subjectivity, our ratings after the first round for each report were in substantial agreement. Furthermore, we had little difficulty in reaching consensus by the third round.

In some cases, the two reports did not provide sufficient details to allow for ratings. To resolve this issue, we contacted the authors of the reports 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.)

At various points in this paper, we cite studies that provide relevant evidence. To ensure that we cited them properly, in December 2007 we sent a copy of our paper to all authors whom we cited in a substantive manner asking them to inform us if we had not properly referred to their findings. None of the authors objected to the way that we had summarized their research. We also invited them to review the paper.

Audit Findings for Amstrup et al. (AMD)

In auditing AMD's forecasting procedures, we first agreed that 24 of the 140 forecasting principles were irrelevant to the forecasting problem they addressed. We then examined principles for which our ratings differed. After two rounds of consultation (i.e., the process involved three rounds in all), we were able to reach consensus on ratings against all 116 relevant principles. We were unable to rate AMD's procedures against 26 relevant principles (Table A.3) due to a lack of information. We attempted to obtain additional information from the authors of the administrative reports, but they refused to cooperate. Full disclosure of the ratings is provided in Tables A.1, A.2, A.3, and A.4 in the Appendix.

Overall, we found that AMD contravened 41 principles and apparently contravened 32. No justifications were provided for any of the contraventions. In all, of the 116 relevant principles, we could only find evidence that AMD properly applied 17 (14.7%) (Table A.4).

We describe some of the more serious problems with the AMD forecasting procedures in the rest of this section of our audit.

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

The forecasts in AMD rely on the opinions of a single polar-bear expert. The opinions were transformed into a complex set of formulae, but were unaided by evidence-based forecasting principles. Unaided experts' judgments are not appropriate for forecasting in this situation as we show in this section.

One of the most counter-intuitive findings in forecasting is that judgmental forecasts by researchers who ignore accepted forecasting principles have little value in complex and uncertain situations (Armstrong, 1978, p. 91-96; Tetlock 2005). The finding applies whether the opinions are expressed in words, spreadsheets, or mathematical models. In relation to the latter, Pilkey and Pilkey-Jarvis (2007) provide examples of the failure of domain experts' mathematical models when applied to diverse natural science problems including fish stocks, beach engineering, and invasive plants. The finding also applies regardless of how much information is used by the experts and whether or not the information is of high quality. Among the reasons for this are:

a) Complexity: People cannot assess complex relationships through unaided observations.
b) Coincidence: People confuse correlation with causation.
c) Feedback: People making judgmental predictions typically do not receive unambiguous feedback they can use to improve their forecasting.

d) Bias: People have difficulty in obtaining or using evidence that contradicts their initial beliefs. This problem is especially serious among people who view themselves as experts.

Despite the lack of validity of unaided forecasts by experts, many public policy decisions are based on such forecasts. 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 the 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."

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 more than 82,000 forecasts. The experts barely if at all outperformed non-experts and neither group did well against simple rules.

Despite the evidence showing that expert forecasts are of no value in complex and uncertain situations, people continue to believe in experts' forecasts The first author's review of empirical research on this problem led him to develop the "Seer-sucker Theory:" "No matter how much evidence exists that seers do not exist, seers will find suckers" (Armstrong 1980).

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 found to be associated with the same outcome, 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."

Regional variability adds to uncertainty. For example, Antarctic ice mass extent has been growing while sea and air temperatures have been increasing (e.g. Zhang 2007). At the same time, depth-averaged oceanic temperatures around the Southeastern Bering Sea (Richter-Menge et al. 2007) have been cooling since 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).

In their abstract, AMD predicted a loss of "…⅔ of the world's current polar bear population by mid-century." The ⅔ figure is at odds with the output from the authors' "deterministic model" as shown in their Table 6. The model's "ensemble mean" predictions are for a much more modest decline in the polar bear population of 17% by the decade centered on 2050. Even the GCM minimum ice scenario used as an extreme input by the authors' provides a forecast decline of 22%: much less than the ⅔ figure in their abstract. As best we can tell, the ⅔ figure was derived informally by the authors' from the outputs of their Bayesian network modeling exercise. The Bayesian network output of interest is in the form of probabilities (expressed as percentages) for each of five possible population states, namely: "larger", "same as now", "smaller", "rare", and "extinct" (Table 8, pp. 66-67). There is, however, no clear link between the sets of probabilities for each population state for each of the authors' four Arctic eco-regions and the dramatic ⅔ population reduction figure.

Despite the uncertainty, instability, and complexity of the situation, AMD made predictions based on assumptions that we view as questionable. They used little historical data. And they preferred extreme forecasts to conservative ones.

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 information from a single source only is used, the validity and reliability of the forecasting process is suspect. 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 from changing sea-ice habitat alone 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 underrated 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 what is predicted during the 21st Century by the GCM projections used by AMD. 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 seem crucial to any forecasts that assume dramatic changes in the bears' environment.

Audit Findings for Hunter et al. (H6)

Hunter et al. (referred to as H6) forecast polar bear numbers and their survival probabilities in the southern Beaufort Sea for the 21st Century.

Of the 140 forecasting principles, we agreed that 35 were irrelevant to the forecasting problem. We then examined principles for which our ratings differed, and after three rounds of consultation we were able to reach consensus on ratings against 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 contravened 61 principles (Appendix Table A.5) and probably contravened an additional 19 principles (Appendix Table A.6). We were unable to rate H6's procedures against 15 relevant principles (Appendix Table A.7) due to a lack of information. Perhaps the best way to summarize H6's efforts is to say that the authors properly applied only ten (9.5%) of the 105 relevant principles (Table A.8).

Many of the contraventions in H6 were similar to those in AMD and we provide the H6 audit details in Tables A.5, A.6, A.7 and A.8 of the Appendix. Here are some examples of contraventions or apparent contraventions, some of which, on their own, raise serious questions about the validity of the H6 forecasts:

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

The H6 authors 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 threat to existing human settlements?

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

H6 relied heavily on five years of data with unknown measurement errors. Furthermore, did the capture data on which they relied provide representative samples of bears in the southern Beaufort Sea given the vast area involved and difficulties in spotting and capturing the bears? Bears wander over long distances and do not respect administrative boundaries (Amstrup et al. 2004). The validity of the data was compromised further by imposing a speculative demographic model on the raw capture-recapture data (Amstrup et al. 2001; Regehr et al. 2006).

Obtain all important data (Principle 4.4)

H6 estimated their key relationship—between ice-free days and the polar bear population—by using data that appear to be unreliable. In addition, they rely on only five years of data with a limited range of climate and ecology combinations. 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., substantially 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).

Given that they used only five observations, the above points are moot. It is impossible to estimate a causal relationship in a complex and uncertain situation by using only five data points.

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

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 (and hence more ice-free days) with lower polar bear survival rates, yet as noted in relation to AMD, cold temperatures have also been found to be associated with the same outcome. Further, and again as noted above in relation to AMD, regional variability, such as sea ice increases while sea and air temperatures go up, add to uncertainty.

There is general agreement that polar bear populations have increased or remained stable in the Alaska regions in recent decades (see e.g., Amstrup et al. 1995; p. 216 of Angliss and Outlaw 2007). In contrast, H6 assumed that there are downward forces that will cause the trend to reverse. Studies in economics have, however, shown little success in predicting turning points. Indeed, Armstrong and Collopy (1993) proposed the principle that one should not extrapolate trends if they are contrary to the direction of the causal forces as judged by domain experts. They tested the principle on four data sets involving 723 long-range forecasts and found that it reduced forecast error by 43%. So, even if one had good reason to expect a trend to reverse, being conservative and avoiding the extrapolation of *any* trend will increase the accuracy of forecasts.

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 (International Union for Conservation of Nature) 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 to be valid nor convincing.

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, and wind conditions. For example Holloway and Sou (2002), Ogi and Wallace (2007), and Nghiem et al. (2007) suggested that large-scale atmospheric winds and related circulatory as well as warming and cooling patterns play an important role in causing—in some situations with significant time delays—both the decline in extent and thinning of Arctic sea ice; those effects were not correctly included in the GCM forecasts of sea ice and hence in the forecasts of the quality of the polar bear habitats.

In addition, Dyck et al. (2007) recently noted that forecasts of polar bear decline 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 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 accepted the GCM model predictions of large declines in summer ice throughout the 21st century, so their forecasts were sensitive to their estimate of the quantitative effect of ice extent on polar bear survival and population growth rates.

Forecast for alternate interventions (Principle 10.7)

H6 did 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 benefits of protecting areas by prohibiting marine and land-based activities at critical times? 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?

Provide easy access to the data (Principle 13.8)

The authors of the reports that we audited did not provide include all of the data they used in their reports. We requested the missing data but they did not provide it.

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

To assess meaningful prediction intervals, 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 the outcome for bears. Extending this exercise to the forecasts of climate and summer ice extent would have further widened the range of possible outcomes.

Discussion

Rather than relying on untested procedures to forecast polar bear populations, the most appropriate approach would be to rely upon prior evidence on which forecasting methods work best in which conditions. By doing this, one can turn to empirical evidence drawn from a wide variety of forecasting problems – such as forecasting human populations.

Given the enormous uncertainty involved in long-term forecasts of polar bear populations, the lack of accurate time-series data on these populations, and the complex relationships that are subject to much uncertainty, prior evidence from forecasting research calls for simple and conservative methods. This means that one should follow a long-term trend if such a trend is consistent and if there are no strong reasons to expect a change in the trend. Lacking a trend, forecasters should turn to the so-called "random walk" or no-change model. The principle of conservatism is sensible, and there is also much empirical evidence to support this notion of "damping trends" (see Armstrong 2001).

An early review of the evidence on how to forecast given uncertainty was provided in Armstrong (1985). For example, Schnaars (1984) and Schnaars and Bavuso (1985) concluded that the random walk was typically the most accurate model in their comparative studies of hundreds of economic series with forecast horizons up to five years. The principle has a long history. For example, regression models "regress" towards a no-change forecasts when the estimates of causal reationships are uncertain.

Summary

We inspected the nine administrative reports that were commissioned by the U.S. government with the stated purpose of supporting the listing of polar bears under the Endangered Species Act. Since the current population of bears is not at a level that causes concern, the case for listing depends upon forecasts of serious declines in bear numbers in decades to come. None of the reports included references to scientific works on forecasting methods.

We found that the two reports most relevant to the listing decision made assumptions where forecasts were required. Even if these assumptions had been valid, the bear population forecasting procedures described in the reports contravened many important forecasting principles. Table 1 summarizes our forecasting audits of the two key reports, Amstrup et al. (2007) and Hunter et al. (2007):

Table 1: Summary ratings of relevant principles from the forecasting audits

Principles	AMD	H6
Contravened	41	61
Apparently contravened	32	19
Not auditable	26	15
Properly applied	17	10
Totals	116	105

Decision makers and the public should require scientific forecasts of both the polar bear population *and* of the costs and benefits of alternative policies before a decision is made on whether to list polar bears as threatened or endangered. We recommend that important forecasting efforts such as this should properly apply all relevant principles and that their procedures be audited to ensure that they do so.

Appendix: Full Disclosure of the Codings

Table A.1: Principles contravened in Amstrup et al. (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 A.2: Principles apparently contravened 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 A.3: Principles not rated 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.

Table A.4: Principles properly applied or *apparently properly applied* in AMD

Setting objectives:

1.1 Describe decisions that might be affected by the forecasts.

Structuring the problem:

2.2 *Tailor the level of data aggregation (or segmentation) to the decisions.*
2.3 *Decompose the problem into parts.*
2.6 *Structure problems that involve causal chains.*

Identify Data Sources:

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

Collecting data:

4.6 *Obtain the most recent data.*

Preparing Data:

5.8 *Use graphical displays for data.*

Selecting Methods:

6.3 Use structured rather than unstructured forecasting methods.
6.5 *Use causal methods rather than naive methods if feasible.*

Implementing Methods: General

7.5 *Adjust for events expected in the future*.
7.6 *Pool similar types of data.*
7.7 *Ensure consistency with forecasts of related series and related time periods.*

Implementing Judgmental Methods:

8.3 *Ask experts to justify their forecasts in writing.*

Implementing Methods: Quantitative Models with Explanatory Variables:

10.7 *Forecast for alternate interventions*.

Presenting Forecasts:

15.2 *Provide complete, simple, and clear explanations of methods.*
15.3 *Describe your assumptions.*

Learning That Will Improve Forecasting Procedures:

16.1 Consider the use of adaptive forecasting models.

Table A.5: Principles contravened in Hunter et al. (H6)

Setting Objectives:

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

Structuring the problem:

2.6 Structure problems that involve causal chains.

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.

Collecting Data:

4.4 Obtain all of the important data

Preparing Data:

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

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.

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.

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.

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.

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.

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.

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.

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).

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 A.6: Principles apparently contravened in H6

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.

Structuring the problem:

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

Identify Data Sources:

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

Collecting Data:

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

Preparing Data:

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

Implementing Methods: General

7.6 Pool similar types of data.

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.

Evaluating Methods:

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

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 A.7: Principles not rated due to lack of information in H6

Setting Objectives:

1.5 Obtain decision makers' agreement on methods

Structuring the problem:

2.7 Decompose time series by level and trend

Identify Data Sources:

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

Collecting Data:

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

Preparing Data:

5.1 Clean the data

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

Evaluating Methods:

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

Presenting Forecasts:

15.3 Describe your assumptions

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

Table A.8: Principles properly applied or *apparently properly applied* in H6

Structuring the problem:

2.2 *Tailor the level of data aggregation (or segmentation) to the decisions.*

Collecting data:

4.6 *Obtain the most recent data.*

Selecting Methods:

6.3 Use structured rather than unstructured forecasting methods.

Implementing Methods: Quantitative Models with Explanatory Variables:

10.1 Rely on theory and domain expertise to select causal (or explanatory) variables.
10.3 *Rely on theory and domain expertise when specifying directions of relationships*.
10.6 *Prepare forecasts for at least two alternative environments.*

Assessing Uncertainty:

14.1 *Estimate prediction intervals (PIs).*
14.2 *Use objective procedures to estimate explicit prediction intervals*.

Presenting Forecasts:

15.4 Present prediction intervals.
15.5 *Present forecasts as scenarios.*

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." The total sum we were paid for our consulting was $9,998. 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 Survey administrative reports 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 Amstrup et al., Hunter et al, or 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 of this paper.

Acknowledgments

We thank Don Esslemont, Milton Freeman, Paul Goodwin, Benny Peiser, Orrin Pilkey, Tom Stewart and two anonymous reviewers for their comments on earlier drafts. Janice Dow and Kelly Jin provided editorial assistance.

References

Amstrup, S.C, Garner, G.W. and Durner, G.M. (1995). Polar Bears in Alaska. In E.T. La Roe (Ed.), Our Living Resources: A report to the nation on the abundance, distributions, and health of the U.S. plants, animals, and ecosystems. Washington, DC: U.S. Department of the Interior-National Biological Sciences, pp 351-353.

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.

Armstrong, J. S. and Collopy, F. (1993). Causal forces: Structuring knowledge for time-series extrapolation. *Journal of Forecasting*, 12, 103-115.

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.

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). Rapid 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.

Pilkey, O. H., and Pilkey-Jarvis, L. (2007). *Useless arithmetic*. New York: Columbia University Press.

Prestrud, P. and Stirling, I. (1994) The international polar bear agreement and the current status of polar bear conservation. *Aquatic Mammals*, 20, 113-124.

Regehr, E.V., Amstrup, S.C. and Stirling, I. (2006). Polar bears population status in the Southern Beaufort Sea. USGS Alaska Science Center, Anchorage, Open File Report 1337.

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.

Schnaars, Stephen P. (1984), "Situational factors affecting forecasting accuracy," *Journal of Marketing Research,* 21, 290-297.

Schnaars, S. P. and Bavuso, R. J. (1986), Extrapolation models on very short-term forecasts. *Journal of Business Research*. 14, 27-36.

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.