



May 2008

## **Polar Bear Population Forecasts: A Public-Policy Forecasting Audit**

J. Scott Armstrong  
*University of Pennsylvania, [armstrong@wharton.upenn.edu](mailto:armstrong@wharton.upenn.edu)*

Kesten C. Green  
*Monash University*

Willie Soon  
*Harvard-Smithsonian Center for Astrophysics*

Follow this and additional works at: [https://repository.upenn.edu/marketing\\_papers](https://repository.upenn.edu/marketing_papers)

---

### **Recommended Citation**

Armstrong, J. S., Green, K. C., & Soon, W. (2008). Polar Bear Population Forecasts: A Public-Policy Forecasting Audit. Retrieved from [https://repository.upenn.edu/marketing\\_papers/134](https://repository.upenn.edu/marketing_papers/134)

Postprint version. To be published in *Interfaces*. Working Paper, Version 77, May 2008.  
URL: [http://www.forecastingprinciples.com/Public\\_Policy/polarbear.html](http://www.forecastingprinciples.com/Public_Policy/polarbear.html)

This paper is posted at ScholarlyCommons. [https://repository.upenn.edu/marketing\\_papers/134](https://repository.upenn.edu/marketing_papers/134)  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

## Polar Bear Population Forecasts: A Public-Policy Forecasting Audit

### Abstract

Calls to list polar bears as a threatened species under the United States Endangered Species Act are based on forecasts of substantial long-term declines in their population. Nine government reports were written to help U.S. Fish and Wildlife Service managers decide whether or not to list polar bears as a threatened species. We assessed these reports based on evidence-based (scientific) forecasting principles. None of the reports referred to sources of scientific forecasting methodology. Of the nine, Amstrup, Marcot, and Douglas (2007) and Hunter et al. (2007) were the most relevant to the listing decision, and we devoted our attention to them. Their forecasting procedures depended on a complex set of assumptions, including the erroneous assumption that general circulation models provide valid forecasts of summer sea ice in the regions that polar bears inhabit. Nevertheless, we 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 during the coming decades. We found that Amstrup et al. properly applied 15 percent of relevant forecasting principles and Hunter et al. 10 percent. Averaging across the two papers, 46 percent of the principles were clearly contravened and 23 percent were apparently contravened. Consequently, their forecasts are unscientific and inconsequential to decision makers. We recommend that researchers apply all relevant principles properly when important public-policy decisions depend on their forecasts.

### Comments

Postprint version. To be published in *Interfaces*. Working Paper, Version 77, May 2008.

URL: [http://www.forecastingprinciples.com/Public\\_Policy/polarbear.html](http://www.forecastingprinciples.com/Public_Policy/polarbear.html)

# **Polar Bear Population Forecasts: A Public-Policy Forecasting Audit**

Version 80: Forthcoming in *Interfaces*

J. Scott Armstrong

The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

[armstrong@wharton.upenn.edu](mailto:armstrong@wharton.upenn.edu)

Kesten C. Green

Business and Economic Forecasting, Monash University, Vic 3800, Australia.

[kestenc@kestengreen.com](mailto:kestenc@kestengreen.com)

Willie Soon

Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138

[wsoon@cfa.harvard.edu](mailto:wsoon@cfa.harvard.edu)

## **Abstract**

Calls to list polar bears as a threatened species under the United States Endangered Species Act are based on forecasts of substantial long-term declines in their population. Nine government reports were written to help U.S. Fish and Wildlife Service managers decide whether or not to list polar bears as a threatened species. We assessed these reports based on evidence-based (scientific) forecasting principles. None of the reports referred to sources of scientific forecasting methodology. Of the nine, Amstrup, Marcot, and Douglas (2007) and Hunter et al. (2007) were the most relevant to the listing decision, and we devoted our attention to them. Their forecasting procedures depended on a complex set of assumptions, including the erroneous assumption that general circulation models provide valid forecasts of summer sea ice in the regions that polar bears inhabit. Nevertheless, we 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 during the coming decades. We found that Amstrup et al. properly applied 15 percent of relevant forecasting principles and Hunter et al. 10 percent. Averaging across the two papers, 46 percent of the principles were clearly contravened and 23 percent were apparently contravened. Consequently, their forecasts are unscientific and inconsequential to decision makers. We recommend that researchers apply *all* relevant principles properly when important public-policy decisions depend on their 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.

Despite widespread agreement that the polar bear population increased during recent years following the imposition of stricter hunting rules (Prestrud and Stirling 1994), new concerns have been expressed that climate change will threaten the survival of some subpopulations in the 21<sup>st</sup> century. Such concerns led the U.S. Fish and Wildlife Service to consider listing polar bears as a threatened species under the United States 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 to list polar bears thus rests on long-term forecasts.

The U.S. Geological Survey commissioned nine administrative reports to satisfy the request of the Secretary of the Interior and the Fish and Wildlife Service to conduct analyses. Our objective was to determine if the forecasts were derived from accepted scientific procedures. We first examined the references in the nine government reports. We then assessed the forecasting procedures described in two of the reports relative to forecasting principles. The forecasting principles that we used are derived from evidence obtained from scientific research that has shown the methods that provide the most accurate forecasts for a given situation and the methods to avoid.

## Scientific Forecasting Procedures

Scientists have studied forecasting since the 1930s; Armstrong (1978, 1985) provides summaries of important findings from the extensive forecasting literature.

In the mid 1990s, Scott Armstrong established the Forecasting Principles Project to summarize 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). Forty internationally recognized forecasting-method experts formulated the principles and 123 reviewed them. We refer to the evidence-based methods as scientific forecasting procedures.

The strongest evidence is derived from empirical studies that compare the performance of alternative methods; the weakest is based on received wisdom about proper procedures. Ideally, performance is assessed by the ability of the selected method to provide useful *ex ante* forecasts. However, some of the 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, were included in the principles list.

The principles were derived from many fields, including demography, economics, engineering, finance, management, medicine, psychology, politics, and weather; this ensured that they encapsulated all relevant evidence and would apply to all types of forecasting problems. Some reviewers of our research have suggested that the principles do not apply to the physical sciences. When we asked them for evidence to support that assertion, we did not receive useful responses. Readers can examine the principles and form their own judgments on this issue. For example, does the principle, “Ensure that information is reliable and that measurement error is low,” not apply when forecasting polar bear numbers?

The forecasting principles are available at [www.forecastingprinciples.com](http://www.forecastingprinciples.com), a website that the International Institute of Forecasters sponsors. The directors of the site claim that it provides “all useful knowledge about forecasting” and invite visitors to submit any missing evidence. The website also provides forecasting audit software that includes a summary of the principles (which currently number 140) and the strength of evidence for each principle; Armstrong (2001) and papers posted on the website provide details.

## **General Assessment of Long-Term Polar Bear Population Forecasts**

We examined all references cited in the nine U.S. Geological Survey Administrative Reports posted on the iInternet at [http://usgs.gov/newsroom/special/polar\\_bears/](http://usgs.gov/newsroom/special/polar_bears/). The reports, which included 444 unique references, were Amstrup, Marcot, and Douglas (2007), Bergen et al (2007), DeWeaver (2007), Durner et al. (2007), Hunter et al. (2007), Obbard et al. (2007), Regehr et al. (2007), Rode, Amstrup, and Regehr (2007), and Stirling et al. (2007). We were unable to find references to evidence that the forecasting methods described in the reports had been validated.

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

We audited the forecasting procedures in the reports that we judged provided the strongest support (i.e., forecasts) for listing polar bears. We selected Amstrup, Marcot, and Douglas (2007), which we will refer to as AMD, because the press had discussed their forecast widely. We selected Hunter et al. (2007), which we will refer to as H6, because the authors used a substantially different approach to the one reported in AMD.

The reports provide forecasts of polar bear populations for 45, 75, and 100 years from the year 2000 and make recommendations with respect to the polar bear-listing decision. However, their recommendations do not follow logically from their research because they only make forecasts of the polar bear population. To make policy recommendations based on forecasts, the following assumptions are necessary:

1. Global warming will occur and will reduce the amount of summer sea ice;
2. Polar bears will not adapt; thus, they will obtain less food than they do now by hunting from the sea-ice platform;
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 that depend on an Endangered Species Act listing.

Regarding the first assumption, both AMD and H6 assumed that general circulation models (GCMs) provide scientifically valid forecasts of global temperature and the extent and thickness of sea ice. AMD stated: “Our future forecasts are based largely on information derived from general circulation model (GCM) projections of the extent and spatiotemporal distribution of sea ice” (p. 2 and Figure 2 on p. 83 of AMD). 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 [Intergovernmental Panel on Climate Change] Fourth Assessment Report (AR4) fully-coupled general circulation models” (p. 11). 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. When they audited the GCM climate modelers’ procedures, they found that only 13 percent of the relevant forecasting principles were followed properly; some contraventions of principles were critical. Their findings were consistent with earlier cautions. For example, Soon et al. (2001) found that the current generation of GCMs is unable to meaningfully calculate the effects that additional atmospheric carbon dioxide has on the climate. This is because of the uncertainty about the past and present climate and ignorance about relevant weather and climate processes. Some climate modelers state that the GCMs do *not* provide forecasts. According to one of the lead authors of the IPCC’s AR4 (Trenberth 2007),

...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.

AMD and H6 provided no scientific evidence to support their assumptions about any of the four issues that we identified above. Thus, their forecasts are of no value to decision makers. Nevertheless, we audited their polar bear-population forecasting procedures to assess if they would have produced valid forecasts *if* the underlying assumptions had been valid.

In conducting our audits, we read AMD and H6 and independently rated the forecasting procedures described in the reports by using the forecasting audit software mentioned above. The rating scale ranged from -2 to +2; the former indicated that the procedures contravene the principle; the latter signified that it is properly applied. Following the initial round of ratings, we examined differences in our ratings to reach consensus. When we had difficulty in reaching consensus, we moved ratings toward “0.” Principle 1.3 (*Make sure forecasts are independent of politics*) is an example of a principle that was contravened in both reports (indeed, in all nine). By politics, we mean any type of organizational bias or pressure. It is not unusual for different stakeholders to prefer particular forecasts; however, if forecasters are influenced by such considerations, forecast accuracy could suffer. The header on the title page of each of the nine reports suggests how the authors interpreted their task: “USGS Science Strategy to Support U.S. Fish and Wildlife Service Polar Bear Listing Decision.” A more neutral statement of purpose might have read “Forecasts of the polar bear population under alternative policy regimes.”

While it was easy to code the two reports’ procedures against Principle 1.3, the ratings were subjective for many principles. Despite the subjectivity, our ratings after the first round of analyses for each report were substantially in agreement. Furthermore, we readily achieved consensus by the third round.

The two reports did not provide sufficient detail to allow us to rate some of the relevant principles. As a result, we contacted the report authors for additional information. We also asked them to review the ratings that we had made and to provide comments. In their replies, the report authors refused to provide any responses to our requests. (See #2 in the *Author comments* section at the end of this paper.)

In December 2007, we sent a draft of this article to all authors whose works we cited substantively and asked them to inform us if we had misinterpreted their findings. None objected to our interpretations. We also invited each author to review our paper, but received no reviews from our requests.

## **Audit Findings for AMD**

In auditing AMD’s forecasting procedures, we first agreed that 24 of the 140 forecasting principles were irrelevant to the forecasting problem they were trying to address. We then examined principles for which our ratings differed. The process involved three rounds of consultation; after two rounds, 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 A3) because the paper lacked the necessary information. Tables A1, A2, A3, and A4 provide full disclosure of our AMD ratings.

Overall, we found that AMD definitely contravened 41 principles and apparently contravened an additional 32 principles. The authors provided no justifications for the contraventions. Of the 116 relevant principles, we could only find evidence that AMD properly applied 17 (14.7 percent) (Table A4).

In the remainder of this section, we will describe some of the more serious problems with the AMD forecasting procedures by listing a selected principle and then explaining how AMD addressed it.

*Principle 6.7: Match the forecasting method(s) to the situation.*

The AMD forecasts rely on the opinions of a single polar bear expert. The report authors transformed these opinions into a complex set of formulae without using evidence-based forecasting principles. In effect the formulae were no more than a codification of the expert's unaided judgments, which are not appropriate for forecasting in this situation.

One of the most counter-intuitive findings in forecasting is that judgmental forecasts by experts who ignore accepted forecasting principles have little value in complex and uncertain situations (Armstrong 1978, p. 91-96; Tetlock 2005). This 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 they are applied to diverse natural science problems including fish stocks, beach engineering, and invasive plants. This finding also applies regardless of the amount and quality of information that the experts use because of the following:

1. Complexity: People cannot assess complex relationships through unaided observations.
2. Coincidence: People confuse correlation with causation.
3. Feedback: People making judgmental predictions typically do not receive unambiguous feedback that they can use to improve their forecasting.
4. 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 expert unaided forecasts, 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 of false expert forecasts, 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 during an Earth day speech at Swarthmore College (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 picked topics (geographic and substantive) both within and outside of their areas of expertise and asked them to forecast the probability that various situations would or would not occur. By 2003, he had accumulated more than 82,000 forecasts. The experts barely, if at all, outperformed non-experts; 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," which states that "No matter how much evidence exists that seers do not exist, seers will find suckers" (Armstrong 1980).

*Principle 7.3: Be conservative in situations of high uncertainty or instability.*

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 uncertain. For example, AMD associated warmer temperatures with lower polar bear survival rates; yet, as the following quote illustrates, colder temperatures have also been found to be associated with the same outcome: “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, Stirling, and Lentfer 1986, p. 249). Stirling (2002, p. 68, 72) further expanded on the complexity of polar bear and 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 has been increasing while sea and air temperatures have also been increasing (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 temperatures by  $1.6 \pm 0.6^\circ\text{C}$ , there was no consistent mid-September (the period of minimal ice extent) ice decline in the Canadian Beaufort Sea over the continental shelf, which had been ice-covered for the 36 years between 1968 and 2003 (Melling, Riedel, and Gedalof 2005).

In their abstract, AMD predicted a loss of “...<sup>2</sup>/<sub>3</sub> of the world’s current polar bear population by mid-century.” The <sup>2</sup>/<sub>3</sub> figure is at odds with the output from the authors’ “deterministic model” as they show in Table 6 in their report. The model’s “ensemble mean” prediction is for a more modest decline of 17 percent in the polar bear population by 2050. Even the GCM minimum ice scenario, which the authors used as an extreme input, provides a forecast decline of 22 percent—much less than the <sup>2</sup>/<sub>3</sub> figure they state in their abstract. We believe that the authors derived their <sup>2</sup>/<sub>3</sub> figure informally 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: “larger,” “same as now,” “smaller,” “rare,” and “extinct” (see Table 8, pp. 66-67 in the AMD report). 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 <sup>2</sup>/<sub>3</sub> population-reduction figure.

AMD made predictions based on assumptions that we view as questionable. They used little historical data and extreme forecasts rather than conservative ones.

*Principle 8.5: Obtain forecasts from heterogeneous experts.*

AMD’s polar bear population forecasts were the product of a single expert. Experts vary in their knowledge and in how they approach problems. A willingness to bring additional information and different approaches to bear on a forecasting problem improves accuracy. When researchers use information from a single source only, the validity and reliability of the forecasting process is suspect. In addition, in situations in which 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).

*Principle 10.2: Use all important variables.*

Dyck et al. (2007) noted that scenarios of polar bear population 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 that polar bears evolved from brown bears 250,000 years ago; however, they appear to have underrated the fact that polar bears probably experienced much warmer conditions in the Arctic over that extended period, including periods in which the sea-ice habitat was less than the amount predicted during the 21<sup>st</sup> century by the GCM projections that AMD used. A 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 was documented by Hamilton and Brigham-Grette (1991), Brigham-Grette and Hopkins (1995), and Norgaard-Pedersen et al. (2007). 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.” The nature and extent of polar bear adaptability seem crucial to any forecasts that assume dramatic changes in the bears’ environment.

## **Audit Findings for H6**

H6 forecast polar bear numbers and their survival probabilities in the southern Beaufort Sea for the 21<sup>st</sup> century.

Of the 140 forecasting principles, we agreed that 35 were irrelevant to the forecasting problem. We found that H6’s procedures clearly contravened 61 principles (Table A5) and probably contravened an additional 19 principles (Table A6). We were unable to rate H6’s procedures against 15 relevant principles (Table A7) because of a lack of information. Perhaps the best way to summarize H6’s efforts is to say that the authors properly applied only 10 (9.5 percent) of the 105 relevant principles (Table A8).

Many of the contraventions in H6 were similar to those in AMD. We describe some of the more serious problems with the H6 forecasting procedures by examining their contraventions of 13 important principles that differed from the contraventions discussed in AMD.

*Principles 1.1–1.3: Decisions, actions, and biases.*

The H6 authors did not describe alternative decisions that might be taken (as Principle 1.1 requires), nor did they propose relationships between possible forecasts and alternative decisions (as Principle 1.2 requires). For example, what decision would be implied by a forecast that predicts that bear numbers will increase to where they become a threat to existing human settlements?

*Principle 4.2: Ensure that information is reliable and that measurement error is low.*

H6 relied heavily on five years of data with unknown measurement errors. Furthermore, we question whether 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, McDonald, and Durner 2004). The validity of the data was also compromised because H6 imposed a speculative demographic model on the raw capture-recapture data (Amstrup, McDonald, and Stirling 2001, Regehr, Amstrup, and Stirling 2006).

*Principle 4.4: Obtain all important data.*

H6 estimated their key relationship—between ice-free days and the polar bear population—by using data that appear to be unreliable primarily because of the difficulty of estimating the polar bear population, but also because of the measurements of ice. Experts in this field, including the authors of the nine reports, are aware of these problems. 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 in Regehr et al. (2007, Figure 3) shows that 1987, 1993, and 1998 were exceptional seasons with more than 150 ice-free 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 well-being (Amstrup, McDonald, and Stirling 2001).

Because 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.

*Principle 7.3: Be conservative in situations of high uncertainty or instability.*

The situation regarding polar bears in the southern Beaufort Sea is complex and uncertain. On the basis of five years of data, H6 associated warmer temperatures (and hence more ice-free days) with lower polar bear survival rates. Yet, as we noted in relation to AMD, cold temperatures have also been found to be associated with the same outcome. In addition, regional variability (e.g., sea ice increases while sea and air temperatures increase) adds to uncertainty.

There is general agreement that polar bear populations have increased or remained stable in the Alaska regions in recent decades (Amstrup, Garner, and Durner 1995, Angliss and Outlaw 2007). H6 assumed that there are downward forces that will cause the trend to reverse. However, studies in economics have 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 percent. Therefore, 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.

*Principle 9.2: Match the model to the underlying phenomena.*

Because of the poor spatial resolution of the GCMs, it is important that readers know the meaning of the “southern Beaufort Sea” (SB) in the H6 report. 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, this assumption by H6 is not valid or convincing.

*Principle 9.5: Update frequently.*

When they estimated their model, H6 did not include data for 2006; the most recent year that was then available. From the supplementary information that Regehr et al. (2007, Figure 3) provide, one finds that the number of ice-free days for the 2006 season was approximately 105—close to the mean of the “good” ice years.

*Principle 10.2: Use all important variables.*

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

Other variables, such as migration, snow, and wind conditions, should also be included. For example, Holloway and Sou (2002), Ogi and Wallace (2007), and Nghiem et al. (2007) suggested that large-scale atmospheric winds and related circulatory and 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. The GCM forecasts of sea ice did not correctly include those effects; hence, the forecasts of the quality of the polar bear habitats also did not.

In addition, as Dyck et al. (2007) noted, forecasts of polar bear decline because of dramatic changes in their environment do not take proper account of the extent and type of polar bear adaptability.

*Principle 10.5: Use different types of data to measure a relationship.*

This principle is important when there is uncertainty about the relationships between causal variables (such as ice extent) and the variable 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 21<sup>st</sup> century; therefore, their forecasts were sensitive to their estimate of the quantitative effect of ice extent on polar bear survival and population growth rates.

*Principle 10.7: Forecast for alternate interventions.*

H6 did not explicitly forecast the effects of different policies. For example, if the polar bear population came under stress because of inadequate summer food, what would be the costs and benefits of protecting areas by prohibiting marine and land-based activities such as tourism, capture for research, and hunting 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?

*Principle 13.8: Provide easy access to the data.*

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

*Principle 14.7: When assessing prediction intervals, list possible outcomes and assess their likelihoods.*

To assess meaningful prediction intervals, it is helpful 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 (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 (e.g., provide wider prediction intervals) about the outcome for the bear population. 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 of which forecasting methods work

best under which conditions. Thus, one could turn to empirical evidence drawn from a wide variety of forecasting problems. This evidence is summarized in the Forecasting Method Selection Tree at <http://forecastingprinciples.com>

Armstrong (1985) provided an early review of the evidence on how to forecast given high uncertainty. Schnaars (1984) and Schnaars and Bavuso (1986) concluded that the random walk was typically the most accurate model in their comparative studies of hundreds of economic series with forecast horizons of up to five years. This principle has a long history. For example, regression models “regress” towards a no-change forecast when the estimates of causal relationships are uncertain.

Because of 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. Therefore, one should follow a trend if such a trend is consistent and if there are no strong reasons to expect a change in the trend. Even then, however, it is wise to dampen the trend towards zero given the increasing uncertainty as the forecast horizon is extended. Empirical evidence supports this notion of “damping trends” (Armstrong 2001). Lacking a trend, forecasters should turn to the so-called “random walk” or no-change model.

Given the upward trend in polar bear numbers over the past few decades, a modest upward trend is likely to continue in the near future because the apparent cause of the trend (hunting restrictions) remains. However, the inconsistent long-term trends in the polar bear population, suggest that it is best to assume no trend in the long-term.

## Summary

We inspected nine administrative reports that the U.S. government commissioned. Because the current polar bear population is not at a level that is causing concern, the case for listing depends upon forecasts of serious declines in bear numbers in future decades. None of these reports included references to scientific works on forecasting methods.

We found that the two reports that we judged most relevant to the listing decision made assumptions rather than forecasts. Even if these assumptions had been valid, the bear population forecasting procedures described in the reports contravened many important forecasting principles. We did forecasting audits of the two key reports (Table 1).

| <u>Principles</u>      | <u>AMD</u> | <u>H6</u> |
|------------------------|------------|-----------|
| Contravened            | 41         | 61        |
| Apparently contravened | 32         | 19        |
| Not auditable          | 26         | 15        |
| Properly applied       | <u>17</u>  | <u>10</u> |
| Totals                 | 116        | 105       |

**Table 1: We summarize our forecasting audit ratings of the AMD and H6 reports against relevant forecasting principles.**

Decision makers and the public should require scientific forecasts of both the polar bear population *and* the costs and benefits of alternative policies before making a decision 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. Failure to apply any principle should be supported by evidence that the principle was not applicable.

## Author Comments

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 received \$9,998 as payment for our consulting. We were impressed by the importance of the issue; therefore, after providing our assessment, we decided to continue work 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; it 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 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 this reply from Steven C. Amstrup on November 30, 2007: “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](http://pubicpolicyforecasting.com) with the auditors’ bios and any information relevant, potential sources of bias.

**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.

Identifying 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.

Identifying 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 to 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  
Because of lack of information in AMD**

|   |  |
|---|--|
| Structuring the problem   | Implementing Methods: Quantitative Models with Explanatory Variables                                 |
| 2.5 Structure problems to deal with important interactions among causal variables.                                  | 10.3 Rely on theory and domain expertise when specifying directions of relationships.                |
| Collecting data   | 10.4 Use theory and domain expertise to estimate or limit the magnitude of relationships.            |
| 4.4 Obtain all of the important data  |  |
| 4.5 Avoid the collection of irrelevant data   |  |
| Preparing Data  | Integrating Judgmental and Quantitative Methods  |
| 5.1 Clean the data.   | 11.1 Use structured procedures to integrate judgmental and quantitative methods.                     |
| 5.2 Use transformations as required by expectations.  | 11.2 Use structured judgment as inputs to quantitative models.                                       |
| 5.3 Adjust intermittent series.   | 11.3 Use pre-specified domain knowledge in selecting, weighting, and modifying quantitative methods. |
| 5.4 Adjust for unsystematic past events.  | 11.4 Limit subjective adjustments of quantitative forecasts.   |
| 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   | Evaluating Methods   |
| 6.6 Select simple methods unless empirical evidence calls for a more complex approach.                              | 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.  |
| Implementing Methods: General   | 13.19 Assess face validity.  |
| 7.2 The forecasting method should provide a realistic representation of the situation                               |  |
| Implementing Judgmental Methods   | Assessing Uncertainty  |
| 8.4 Provide numerical scales with several categories for experts' answers.  | 14.12 Do not assess uncertainty in a traditional (unstructured) group meeting.                       |
|   | Learning to 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 (italics)* 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.*

Identifying 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 to Improve Forecasting Procedures

16.1 Consider the use of adaptive forecasting models.

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

|  |  |
|--|--|
| Setting Objectives   | Implementing Quantitative Methods  |
| 1.3 Make sure forecasts are independent of politics.   | 9.1 Tailor the forecasting model to the horizon.   |
| 1.4 Consider whether the events or series can be forecasted.   | 9.2 Match the model to the underlying phenomena.   |
|  | 9.3 Do not use “fit” to develop the model.   |
|  | 9.5 Update models frequently.  |
| Structuring the problem  |  |
| 2.6 Structure problems that involve causal chains.   | Implementing Methods: Quantitative Models with Explanatory Variables   |
|  | 10.2 Use all important variables.  |
| Identifying Data Sources   | 10.5 Use different types of data to measure a relationship.  |
| 3.4 Use diverse sources of data.   | 10.7 Forecast for alternate interventions.   |
| 3.5 Obtain information from similar (analogous) series or cases. Such information may help to estimate trends. | 10.9 Shrink the forecasts of change if there is high uncertainty for predictions of the explanatory variables. |
| Collecting Data  |  |
| 4.4 Obtain all of the important data   | Integrating Judgmental and Quantitative Methods  |
| <i>Preparing Data:</i>   | 11.1 Use structured procedures to integrate judgmental and quantitative methods.                               |
| 5.2 Use transformations as required by expectations.   | 11.2 Use structured judgment as inputs to quantitative models.   |
| 5.4 Adjust for unsystematic past events.   | 11.3 Use pre-specified domain knowledge in selecting, weighting, and modifying quantitative methods.           |
| 5.5 Adjust for systematic events.  |  |
| Selecting Methods  | Combining Forecasts  |
| 6.1 List all the important selection criteria before evaluating methods.                                       | 12.1 Combine forecasts from approaches that differ.  |
| 6.2 Ask unbiased experts to rate potential methods.  | 12.2 Use many approaches (or forecasters), preferably at least five.   |
| 6.6 Select simple methods unless empirical evidence calls for a more complex approach.                         | 12.3 Use formal procedures to combine forecasts.   |
| 6.7 Match the forecasting method(s) to the situation.  | 12.8 Combine forecasts when there is uncertainty about which method is best.                                   |
| 6.8 Compare track records of various forecasting methods.  | 12.9 Combine forecasts when you are uncertain about the situation.   |
| 6.10 Examine the value of alternative forecasting methods.   | 12.10 Combine forecasts when it is important to avoid large errors.  |
| Implementing Methods: General  | Evaluating Methods   |
| 7.1 Keep forecasting methods simple.   | 13.1 Compare reasonable methods.   |
| 7.2 The forecasting method should provide a realistic representation of the situation.                         | 13.2 Use objective tests of assumptions.   |
| 7.3 Be conservative in situations of high uncertainty or instability.  | 13.3 Design test situations to match the forecasting problem.  |
| 7.4 Do not forecast cycles.  | 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.

*Identifying 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

*Identifying 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 to 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.*

## Acknowledgments

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

## References

- Amstrup, S. C., G. W. Garner, G. M. Durner. 1995. Polar Bears in Alaska. 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. U.S. Department of the Interior-National Biological Sciences, Washington, D.C. 351–353.
- Amstrup, S. C., B. G. Marcot, D. C. Douglas. 2007. Forecasting the rangewide status of polar bears at selected times in the 21st century. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Amstrup, S. C., T. L. McDonald, G. M. Durner. 2004. Using satellite radiotelemetry data to delineate and manage wildlife populations. *Wildlife Soc. Bull.* **32** 661–679.
- Amstrup, S. C., T. L. McDonald, I. Stirling. 2001. Polar bears in the Beaufort Sea: A 30-year mark-recapture case history. *J. Agricultural Biol. Environ. Statist.* **6** 221–234.
- Amstrup, S. C., I. Stirling, J. W. Lentfer. 1986. Past and present status of polar bears in Alaska. *Wildlife Soc. Bull.* **14** 241–254.
- Angliss, R. P., R. B. Outlaw. 2007. Alaska marine mammal stock assessments, 2006. Retrieved June 4, 2008, <http://www.nmfs.noaa.gov/pr/pdfs/sars/ak2006.pdf>.
- Armstrong, J. S. 1978. *Long-Range Forecasting: From Crystal Ball to Computer*. Wiley-Interscience, New York, NY.
- Armstrong, J. S. 1980. The Seer-sucker theory: The value of experts in forecasting. *Tech. Rev.* **83** 16–24 (June-July).
- Armstrong, J. S. 1985. *Long-Range Forecasting: From Crystal Ball to Computer, 2<sup>nd</sup> ed.* Wiley-Interscience, New York, NY.
- Armstrong, J. S. 2001. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Kluwer Academic Publishers, Norwell, MA.
- Armstrong, J. S., F. Collopy. 1993. Causal forces: Structuring knowledge for time-series extrapolation. *J. Forecasting* **12** 103–115.
- Bergen, S., G. M. Durner, D. C. Douglas, S. C. Amstrup. 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. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Brigham-Grette, J., D. M. Hopkins. 1995. Emergent marine record and paleoclimate of the last interglaciation along the Northwest Alaskan coast. *Quaternary Res.* **43** 159–173.
- Cerf, C., V. Navasky. 1998. *The Experts Speak*. Random House, New York, NY.
- DeWeaver, E. 2007. Uncertainty in climate model projections of Arctic sea ice decline: An evaluation relevant to polar bears. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Durner, G. M., D. C. Douglas, R. M. Nielson, S. C. Amstrup, T. L. McDonald. 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. Administrative Report. USGS Alaska Science Center, Anchorage, AK.
- Dyck, M. G., S. Romberg. 2007. Observations of a wild polar bear (*Ursus maritimus*) successfully fishing Arctic charr (*Salvelinus alpinus*) and Fourhorn sculpin (*Myoxocephalus quadricornis*). *Polar Biol.* **30** 1625–1628.

- Dyck, M.G., W. Soon, R. K. Baydack, D. R. Legates, S. Baliunas, T. F. Ball, L. O. Hancock. 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., J. S. Armstrong. 2007. Global warming: Forecasts by scientists versus scientific forecasts. *Energy and Environ.* **18** 997–1021.
- Hamilton, T. D., J. Brigham-Grette. 1991. The last interglaciation in Alaska: Stratigraphy and paleoecology of potential sites. *Quaternary Internat.* 10–12, 49–71.
- Holloway, G., T. Sou. 2002. Has Arctic sea ice rapidly thinned? *J. of Climate* **15** 1691–1701.
- Hunter, C. M., H. Caswell, M. C. Runge, S. C. Amstrup, E. V. Regehr, I. Stirling. 2007. Polar bears in the Southern Beaufort Sea II: Demography and population growth in relation to sea ice conditions. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Melling, H., D. A. Riedel, Z. Gedalof. 2005. Trends in the draft and extent of seasonal pack ice, Canadian Beaufort Sea. *Geophysical Res. Lett.* **32**(24) L24501, doi:10.1029/2005GL024483.
- Nghiem, S.V., I. G. Rigor, D. K. Perovich, P. Clemente-Colon, J. W. Weatherly, G. Neumann. 2007. Rapid reduction of Arctic perennial sea ice. *Geophysical Res. Lett.* **34** L19504, doi:10.1029/2007GL031138.
- Norgaard-Pedersen, N., N. Mikkelsen, S. J. Lassen, Y. Kristoffersen, E. Sheldon. 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., T. L. McDonald, E. J. Howe, E. V. Regehr, E. S. Richardson. 2007. Trends in abundance and survival for polar bears from Southern Hudson Bay, Canada, 1984-2005. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Ogi, M., J. M. Wallace. 2007. Summer minimum Arctic sea ice extent and the associated summer atmospheric circulation. *Geophysical Res. Lett.* **34** L12705, doi:10.1029/2007GL029897.
- Pilkey, O. H., L. Pilkey-Jarvis. 2007. *Useless Arithmetic*. Columbia University Press, New York.
- Prestrud, P., I. Stirling. 1994. The international polar bear agreement and the current status of polar bear conservation. *Aquatic Mammals* **20** 113–124.
- Regehr, E. V., S. C. Amstrup, I. Stirling. 2006. Polar bear population status in the Southern Beaufort Sea. <http://pubs.usgs.gov/of/2006/1337/pdf/ofr20061337.pdf>. Retrieved November 17, 2006..
- Regehr, E.V., C.M. Hunter, H. Caswell, S. C. Amstrup, I. Stirling. 2007. Polar bears in the Southern Beaufort Sea I: Survival and breeding in relation to sea ice conditions, 2001–2006. Administrative Report, USGS Alaska Science Center, Anchorage, AK. .
- Richter-Menge, J., J. Overland, A. Proshutinsky, V. Romanovsky, R. Armstrong, J. Morison, S. Nghiem, et al. 2007. State of the climate in 2006: Arctic. *Bull. Amer. Meteorological Soc.* **88** S62–S71.
- Rode, K. D., S. C. Amstrup, E. V. Regehr. 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. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Schnaars, S. P. 1984. Situational factors affecting forecasting accuracy. *J. Marketing Res.* **21** 290–297.
- Schnaars, S. P., R. J. Bavuso. 1986. Extrapolation models on very short-term forecasts. *J. Bus. Res.* **14** 27–36.
- Soon, W., S. Baliunas, S. B. Idso, K.Y. Kondratyev, E. S. Posmentier. 2001. Modeling climatic effects of anthropogenic carbon dioxide emissions: Unknowns and uncertainties. *Climate Res.* **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., T. L. McDonald, E. S. Richardson, E. V. Regehr. 2007. Polar bear population status in the Northern Beaufort Sea. Administrative Report, USGS Alaska Science Center, Anchorage, AK.
- Tetlock, P. E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press, Princeton, NJ.
- Trenberth, K. 2007. Predictions of climate. Retrieved June 2, 2008.  
[http://blogs.nature.com/climatefeedback/2007/06/predictions\\_of\\_climate.html](http://blogs.nature.com/climatefeedback/2007/06/predictions_of_climate.html).
- Zhang, J. 2007. Increasing Antarctic sea ice under warming atmospheric and oceanic conditions. *J. Climate* **20** 2515–2529.