Uncertainties in climate prediction[1]

Gerbrand Komen

When discussing future climate, we reason about the future from the past. This is an example of inductive reasoning, a topic addressed by many philosophers. Below I argue how their arguments might be useful, both for climate researchers and for policy makers.

What is induction?

Induction is a form of reasoning that makes generalizations based on individual instances. Inductive reasoning is the process of reasoning in which the premises of an argument are believed to support the conclusion but do not entail it; i.e. they do not ensure its truth. The philosophical question is whether inductive reasoning is valid. The Scottish philosopher David Hume noted that our everyday reasoning depends on patterns of repeated experience rather than deductively valid arguments. For example, we believe that bread will nourish us because it has done so in the past, but this is not a guarantee that it will always do so. As Hume said, someone who insisted on sound deductive justifications for everything would starve to death. Instead of approaching everything with unproductive skepticism, Hume advocated a practical skepticism based on common sense, where the inevitability of induction is accepted. In the 20th century, thinkers such as Karl Popper and David Miller have disputed the necessity and validity of any inductive reasoning.

So far, just a few quotes from Wikipedia. Personally, I very much like the work of E.T. Jaynes. His monograph “Probability Theory”, which appeared in 2003, has the impressive subtitle “The logic of Science”. Jaynes argues that there are degrees of probability that can be quantified using the Bayesian paradigm, and he gives many examples of how one can combine existing knowledge with new information to obtain “better” knowledge.

Nothing is certain

My own understanding is as follows. Nothing is ever absolutely certain. Some regularities have been observed so often that we find it hard to believe that they might be refuted. Examples are the basic laws of nature within their domain of validity. Think of Newton‟s second law. You can safely predict that this law will also hold tomorrow, even if it is not 100 % certain. This is different when we make statements about the evolution of complex systems. For example, you plan a trip, and expect that a train will appear at a particular place at a particular time with a particular destination. From experience you know that the train may not appear. In fact you may have some statistical information on the reliability of the train company. When the statistics is not available you still may have a prejudice: Japanese trains are always in time, in other countries this may not always be the case. But even in Japan trains may be late, and in other countries they may be surprisingly punctual.

After I joined the Royal Netherlands Meteorological Institute in 1977, I was put in charge of the numerical ocean wave model. There was already a model up and running, so one of the first things I did was to set up a routine verification procedure. Four times a day we collected wave observations

[1] 1 Published on 7 July 2008 as Guest Weblog on „Climate Science: Roger Pielke Sr. Research Group News‟. (http://pielkeclimatesci.wordpress.com/)

from several locations and compared them with the model results, both the analysis and the forecasts. After a few years we had a useful database. Our verification focused on bias and standard deviation for a few key parameters. We calculated monthly values. It was interesting to see how these monthly values varied from month to month, often for no apparent reason. Of course, we tried to improve the model. In fact, we replaced the model by a new model with better physics. Again we did a lot of validation, and we knew the model could skillfully predict the sea state for a couple of days ahead. But even for that model surprises in performance occurred. For example, the model tended to underestimate the wave height at the height of a storm, and it was not always clear whether we should blame too low input winds or something missing in the model physics. [Nevertheless, the model products were deemed useful. The model became operational at ECMWF and many other centres, and the recently completed Global Wave Atlas is consulted frequently.]

Let me summarize. No statement about the future is necessarily (logically) true. But we consider some statements as more likely than others. We derive this likelihood from past experience. Even if this past experience comes in the form of a quantitative measure, like the skill of our wave model, there is no guarantee that a particular prediction will come true. (I believe this also holds in the case of ensemble prediction when we predict pdfs).

What does this mean for climate prediction?
We know it is difficult to predict climate because of the complexity of the socio-economic/biogeophysical system. There are many multiple feedbacks requiring quantitative treatment. This is most realistically done in numerical models, which however suffer from 1. internal variability with limited predictability; 2. the problems of subgrid parametrization; and 3. unresolved processes, either not treated or treated as external forcings. The obvious question is then how serious model predictions should be taken. An important measure is the model skill. However, there is a basic problem: despite some recent progress on establishing correlation skill in predicting observed ten-year mean surface temperature anomalies a decade in advance (Keenlyslide et al, Nature 453, p84-87), it is virtually impossible to test skill for longer time scales. More importantly, skill in the past does not guarantee skill in the future. Therefore, we should also consider other measures of reliability, such as successful hindcasts and the ability of a model to represent key processes. Intercomparison of independent models can also provide valuable information (see e.g., http://www.clivar.org/organization/aamp/publications/mips.htm), the amount of disagreement being a (not the) measure of uncertainty. None of these measures (including skill) guarantee the correctness of the predictions, but it is obvious that some model predictions should be given more confidence than others.

How do you quantify confidence?
Climate scientists are aware of these grades of uncertainty. Those who contributed to the Fourth Assessment Report of IPCC have quantified this by providing a “consistent and transparent treatment of uncertainties”. This resulted in statements of the type “Continued greenhouse gas emissions at or above current rates would cause further warming and induce many changes . . that would very likely be larger than those observed in the 20th century.” Another example is “the climate sensitivity is likely to be in the range 2 °C to 4.5 °C . . “. „Likely‟ and „very likely‟ are defined as > 66 % probability and > 90 % probability, respectively. I think the authors did a wonderful job in trying to quantify (un)certainty. However, it is important to note that the values given are assessed best estimates, based on expert judgment. Let me use an (over)simplified example to illustrate what this means.

Suppose, there are 5 models in which the experts have equal confidence; 4 of these models predict A and the other model predicts B. In that case the assessed likelihood of A being true would be 80 %. There is nothing wrong with this (it is in fact very much along the lines of E.J. Jaynes‟ treatment of probability), but another expert group might come up with another estimate, for example when they consider one of the models substandard. Therefore, this type of probability has a clear subjective element, and differs from the frequentist (Kolmogorov) probability concept I was brought up with. IPCC avoids the word subjective, but dutifully states that uncertainty ranges are assessed, and often based on expert judgment. Unfortunately, this is not properly appreciated by everyone. One scientist told me she thought the likelihoods were objective, because they were expressed as numerical values. And a policy maker told me that she considered the probabilities objective because they came from scientists.

What does this imply?

For decision makers things may become more complicated when they are better aware of the uncertainties and their nature. Decision making is deciding on a line of action, when, nearly always, understanding is limited. In these situations human values, such as attitudes to risk and the extent to which expert judgment is taken into account also play a role.

For scientists the implications are marginal. After all, dealing with uncertainty is their core business. However, scientists should acknowledge that skill is not the only measure for judging model predictions. More importantly, they should make an effort to better communicate the limitations of model predictions, and, more in general, the uncertainties in our understanding of climate change. In The Netherlands we have made an effort to do so, with some success, and with some frustration. IPCC did a decent job in Chapter 6 (Robust Findings, Key Uncertainties) of the Synthesis Report , a chapter deserving better publicity. Otherwise it is business as usual: collect observations, analyse them, and improve models and understanding. Of course, priorities have to be set. I believe this is best done in a transparent discussion avoiding the pitfalls of groupthink (see e.g. Lee Smolin‟s The Trouble with Physics for more on this) and with an open eye for the problems of our planet.