AN EXPERT-OPINION APPROACH TO THE PREDICTION PROBLEM IN COMPLEX SYSTEMS

GERBRAND J. KOMEN

Royal Netherlands Meteorological Institute (KNMI)
P.O. Box 201, 3730 AE De Bilt, The Netherlands
e-mail: komen@knmi.nl

Summary

The use of model forecasts for decision making should be optimized. With this in mind, the concept of modelling the future is discussed from an epistemological point of view and on the basis of a stochastic model interpretation. Traditional definitions of model statistics make reference to an ensemble of systems. Since this does not work for a complex system with a unique state, an alternative approach, based on the subjective (Delphi) opinion of a group of experts, is also considered. This approach is then generalized to the situation in which a set of competing models is available. With a Delphi method a certain likelihood can be assigned to each model. Once the statistics is defined, one may face the issue of predictability. In hindsight (in a 'hindcasting mode') models can be validated by checking how accurate they have been describing observations and they can be falsified when their predictions differ in an unlikely way from the observations. 'Forecasting' is different, because models can never be proven. Therefore, exact prediction of the future is impossible. Definitions of predictability (two examples will be given) necessarily refer to the range of modelled possibilities. It is argued that all model predictions - also those resulting from physical models - should be considered as scenarios. To make rational decisions the likelihood of all possible model forecasts has to be taken into account. In case of complex systems and difficult decisions it appears useful to consider a large variety of models. Experts need not strive for consensus, because a diversity of opinions could lead to better decisions. It is recommended that more attention is paid to Delphi aspects of forecast likelihoods.

Keywords Predictability, complex models, likelihood, expert-opinion.

1. Motivation

The present discussion - although rather general and applicable to many different systems - was inspired by questions related with climate modelling. As is well known the atmospheric concentration of CO_2 and other Greenhouse gases increases due to human activities. It is expected that this will lead to a disturbance of the natural climate. This has led to political discussions, and to an increase of interest in (numerical) climate models.

In the development of these models two trends can be seen. The most advanced physical models of the coupled atmosphere-ocean system are still inadequate in describing (details of) the present climate. Therefore, one seeks improvement to obtain better description of the actual climate and more reliable 'predictions' of climate change. Improvements are expected to come from an enhancement of the spatial resolution, and from the use of more realistic sub-grid scale parametrizations, such as describing for example the cloud-radiation interaction or the effect of ocean waves on air/sea exchange. An example of this approach is given by Washington and Meehl (1989). A summary of similar approaches is given in the IPCC report (1990, 1992).

However, it is realized that chemical, biological and socio-economical factors are also crucial for a correct description of the anthropogenic effects on climate. To model these, simple physical models have been coupled with chemical, biological and socio-economic models (Rotmans, 1990).

Opponents of these latter models criticize by pointing out that the physical subsystem is modelled rather inaccurately, whereas the accuracy of the other subsystems is even less well known. Proponents argue that the interaction of the physical subsystem with the rest cannot be ignored.

This note tries to sort out the underlying assumptions, in an attempt to take away the prevailing confusion. The ideas are not new, but it is hoped that presenting them here may help stimulate the discussion.

Forecast models are often interpreted as stochastic models, predicting probabilities. Therefore, we begin with a discussion of these models. Next we will discuss the concept of probability, which is essential for understanding predictability and the meaning of model forecasts.

2. Stochastic models

Consider first a closed model system, which can be described by n prognostic variables X, and *for which we know the law of evolution* M apart from the value of (a set of) parameters and forcing variables A:

$$M(A)\colon X(t_0) \to X(t). \tag{1}$$

The evolution operator M is nonlinear, in general, and acts on the state vector X to compute the state at a later time. In physical models it usually results from the discretization of a set of (integro-)differential equations. To be specific X may be thought of as the positions and velocities of interacting point particles; A would be their masses and M would be given by classical mechanics.

In this approach X and A are random variables (X is a 'stochastic process', see for example, Doob, 1953), so they define probability distributions f_X and f_A with the property that

$$f_X(x)dx \tag{2}$$

is the probability that X has a value between x and $x + dx$, and similarly for f_A.

In practice, we are often dealing with very complex systems with many degrees of freedom *for which we have only limited knowledge about the laws of evolution*. An example is an atmospheric model, in which case X would represent air pressure, velocity, density, temperature, etc. of the atmosphere on a finite difference grid on the globe at specified levels. But X can also be much larger: it could include ocean and sea-ice variables, the chemical composition of the atmosphere, emission rates, oil price, inflation rate, deforestation rate, etcetera. Typically X has $10^6 - 10^8$ components or more.

One of the central problems in modelling is the correct choice of the (high dimensional vector) space S, in which to describe the phenomena of interest. The choice of parameters

A and dynamical variables X requires a careful analysis of the system and the objectives of the study. In one approach the CO_2 emission rate could be prescribed as a parameter; in another model it could be treated as a dynamical variable depending on energy price, population growth and other variables. Often, the distinction between dynamic variables and parameters is somewhat artificial. Therefore, it is interesting to compare the performance of models in different spaces S. This is most easily formulated by realizing that (1), for a given choice of dynamical variables, can also be interpreted as a collection of mappings (models) $\{M_a\}$ labelled with the possible values a of A, each with their given probability distribution. The generalization to the case of variable size of X and A is obtained by considering the collection of (all) proposed models $\{S_\alpha, M_\alpha\}$ specified by the mappings

$$M_\alpha : X(t_0) \rightarrow X(t), \;\; X \in S_\alpha \,. \tag{3}$$

This collection comprises (1), but it is more general because it simultaneously includes models in which the system is represented by state vectors in different spaces S_α. The label α is not a scalar. It has one component labelling the different models (i.e. sets of differential equations); the other components label the possible values of A. One may think of atmospheric grid point models with different spatial (and temporal) resolution but also of more complex biosphere models which do or do not include certain variables that may or may not be marginal for understanding the system (blue algae or dimethylsulfide concentration, for example).

3. The definition of probability

The definition of probability comes with a conceptual difficulty. Traditionally, one considers the abstract concept of an ensemble of realizations of the system, each with different values for the random components. When predicting the trajectory of a billiard-ball one may specify the uncertainty in its initial position by performing many independent measurements. This defines the ensemble - many billiard-tables with the balls initially in slightly different positions - and the required probability distribution. In weather prediction the distribution of A of (1) can be obtained by considering many *similar* situations in the past, and by determining the corresponding realizations a of A.

However, this procedure cannot be followed in very complex models of the world, because for all we know our world is quite unique at a given time, and therefore it is simply not possible to find *similar* situations. In these cases it is still possible to use statistical concepts, albeit on a rather different basis.

First, for simplicity, consider models of the type (1) and assume that the initial conditions are accurately known, so that the uncertainty is in A. To obtain the desired probability distribution one could then organize a simple Delphi-like procedure (see remark below) in which the opinions of experts are used. This can be done in infinitely many ways, but to be specific consider the case where each expert is asked to give the value of a that he finds most likely. From this then follows a probability distribution of A, not based on a direct objective analysis of past events, but expressing the subjective assessment of the expert panel. For a given panel this specifies the model uniquely. Of course, the width of the distribution of A contains interesting information, a narrow distribution indicating consensus, a broad one meaning disagreement. Also, one could form different panels and compare the results along the rules of test theory (see for example, Crocker and Algine, 1986).

The same procedure may be applied to models of the type (3). Again there is an infinity of possible methods from which one must select and again the specific questions asked to the experts would be an integral part of the definition. An example? Suppose one considers six 'models', consisting of two different (sets of) differential equations, each with one undetermined constant which may assume three different values. Then one can ask the experts to rank them in order of likelihood of describing correctly the change in mean global sea level in the second half of the 21st century. The probabilities so defined specify the likelihood of the models. They are a (subjective) measure expressing the confidence experts have in the correctness of a particular model out of a given set. In this case it is even more interesting than before to know the width of the distribution and to apply test theory to the results of different expert groups.

In my opinion the traditional physical approach should be used whenever possible, in particular for the modelling of physical subsystems, because the results of this approach are much more likely to be correct. The gravitational acceleration at the surface of the earth will be approximately 9.8 m/s^2 also in 2050. However, to estimate the interest rates in that year the Delphi procedure could be helpful, because it quantifies the subjective estimate of the uncertainty in this parameter and allows for numerical modelling of the consequences of its uncertainty.

Remark In its original meaning the name Delphi method (see e.g. Linstone and Turoff, 1975) refers to a technique developed for obtaining judgements from a group of experts. Characteristics are feedback, anonimity and statistical presentation. Here we use the term in a loose way. Most conventional applications strive for consensus. The present application leaves room for feedback, but consensus is not necessary.

4. The use of models

The models can be used in two ways *independently from how the input probabilities have been defined.*

In the first type of model application ('hindcasting') one simulates the past and compares model results with observations. These observations - let us denote them by O - are themselves stochastic variables because of measurement errors. The so-called model counterparts of the observations O^m are functions of the model state vector X. Model I is usually considered to be better than model II in describing a particular observation O_i when the mean and variance (one must choose how to weight) of a time (or spatial) series $O_i - O_i^m(X)$ are smaller for I than for II. If the difference between O_i and $O_i^m(X)$ is unlikely large the model is falsified. One may then attempt to construct a better model and, in fact, it is along these lines that the understanding of the system is enlarged. In practice, one usually compares time series of particular realizations of O and O^m. In true stochastic models one could also attempt to validate the predicted probability distributions by comparing them with the distribution of observations obtained by averaging over analogous situations, but this is not usually done, and, in fact, as discussed in section 3, this would be impossible in a unique complex system.

In the second type of application ('forecasting') one forecasts the future. The justification is most easily expressed for traditional deterministic systems: (A) a system that evolves according to model M will lead to a state $x(t)$; (B) suppose that reality behaves as such a system; then (C) in reality we expect state $x(t)$ to occur at time t. Obviously, (B) is an assumption and for this reason at the moment the prediction is made there is no way of establishing its future correctness. One may hit a billiard-ball in the right direction, expecting with 100% certainty that the desired collision will occur, but the actual collision may never come, because unexpectedly the billiard-table collapses due to woodworm.

For a stochastic model the argument would go as follows: (A) a system that evolves according to the set of models $\{M_\alpha\}$ will lead to a state $X(t)$ (a set of states x_α each with a certain likelihood); (B) suppose that reality behaves as such a system; then (C) in reality we may expect to find state $x_\alpha(t)$ at time t with corresponding likelihood. (The meaning of this expectation in unique complex systems would differ from the conventional physical meaning). But, again, at the moment the prediction is made there is no way of establishing its future correctness. Our sophisticated stochastic climate predictions may be jeopardized by the unexpected penetration of a large meteorite through the ocean bottom.

5. Predictability

Let us first consider models of the type (1). We cannot, in general, compute a definite value for $X(t)$, because neither $X(t_0)$ nor A is known exactly. However, on the basis of statistical information on $X(t_0)$ and A, we can make a statement about the probability of finding a certain realization $x(t)$. Given the probability characteristics of A and X the probability distribution of $X(t)$ follows:

$$M\colon\ (f_A(a),\ f_{X(t_0)}(x)) \ \rightarrow f_{X(t)}(x)\ . \tag{4}$$

Equation (3) is realized most easily with the help of a Monte Carlo simulation. In such a simulation one generates values for A and $X(t_0)$ on the basis of their probability distributions. For each set a, $x(t_0)$ one solves (1) which then automatically generates the probability distribution of $X(t)$.

In this context predictability can be defined as the inverse width of the probability distribution of $X(t)$, or more precisely, for a particular observable O_i, as the inverse width of $f_{O_i(X_t)}(o_i)$. In chaotic systems this width gets quickly very large, even if the initial condition and the values A are known rather accurately. With this definition of predictability, it is quite possible for certain observables to be more predictable than others. Note that this concept of predictability necessarily refers to the range of *modelled* possibilities.

For models of the type (3) the same arguments can be given. Once a likelihood has beens assigned to each M_α, it is, in principle, straightforward to compute the corresponding likelihood distribution of a predicted observable O_i^m.

It is fascinating to speculate about something more profound. To this end consider a hierarchy of stochastic models $\{S_n, M_n\}$ (now labelled with a single integer n), more and more refined with ever more dynamical variables. One might order them according to complexity, say

$$\text{dimension } S_{n_1} \geq \text{dimension } S_{n_2} \quad \text{if} \quad n_1 > n_2 \tag{5}$$

Each model will predict a random state $X_n(t)$ from which the corresponding probability distribution of some observable $O_i^m[X_n(t)]$ is readily computed. We can then consider a sequence of model predictions $\{O_i^n[X_n(t)]\}$ and we would call an observable predictable if

$$\lim_{n \to n_{max}} O_i^n(X_n) \tag{6}$$

exists, with n_{max} the label of the most complex model considered. The basic idea is that some observables are sensitive to additional complexities of the system, whereas others are stable when you make the model more complex. To my knowledge such sequences have not been studied. I expect that for a given sequence of models some observables will be predictable and others will not, whereas for each observable one can construct a set of models in which this observable is not predictable. I have not attempted to prove these conjectures. They seem to be related to our inability to know the future and could perhaps explain why some observables are much harder to predict than others.

6. The psychology of decision making

In daily life, as in physics, one always makes 'models' of the future. These do not need to take the form (1) or (3), but they have in common with (1) and (3) that they are representations of the system in the neural network of the human brain, and that their predictions of the future need not come true. For instance, you want to go out and you expect that it will rain, so you pick up an umbrella. But then the model turns out to be inadequate, because there may be so much wind that you cannot use the umbrella or another unexpected event (not 'modelled') may prevent you from leaving the house.

Often people make decisions with a more or less explicit concept of the desired future situation in mind. They then select from possible courses of action by using (numerical) models to estimate the effect of these decisions on the future. From the foregoing discussion it will be clear that these models do not 'know' the future, but they can be used to generate possible future states with a certain likelihood. The argument then goes: if we do this, then there is a probability that such and so. After scanning all possible decisions they make that decision that brings them with the highest likelihood closest to the desired situation (see for example Lindley, 1985). It is interesting to note that decisions are always made on the basis of models, never on the basis of knowledge of the future state.

7. Conclusions

- A Delphi approach makes it possible to use stochastic mathematical models for forecasting, even for a complex system in which certain parameters are not known from experience.

- We have seen that it is not possible to base decisions on *knowledge* of the future. In fact, one always uses 'models' to generate scenarios with an attached likelihood. If the model predictions do not come true, decisions may have taken us away from the desired state, rather than that they have brought us closer to it.

- Physical models of simple, (nearly) closed systems have a very large likelihood.

- In case of complex systems, and difficult decisions, it appears useful to consider a large variety of models.
- Experts need not strive for consensus. A diversity of opinions could lead to better decision making, because it allows one to consider a larger number of possible futures, which reduces the risk of overlooking an important scenario with a small likelihood (a bifurcation in the thermohaline circulation of the world ocean, for example).
- It might be useful to pay more attention to details of Delphi procedures, such as the selection of (groups of) experts and the formulation of the questions asked.

8. Concluding remark

This note sketched a rationalistic approach to decision making, which pretends that one can control the system, to some extent at least. Of course, there are other approaches as well. For example, some people - they may be called ethicists - act according to certain principles ('thou shalt not lie') which they follow, irrespective of where it takes them. True modellers can not be stopped by this. On the contrary, they will be inspired to add two kinds of human actors to their models - rationalists and ethicists - and they would attempt to model the consequences of the interaction between these actors and the rest of the system.

Acknowledgements

My interest in the subject was stimulated by discussions with Henk Tennekes, who read a draft of this paper, as did Arie Kattenberg, Theo Opsteegh and Wieger Fransen. I would like to thank them all, for their suggestions and encouragement. The idea of the expert-opinion approach to likelihood came up during a discussion with Klaus Hasselmann, when we drove from Tallinn to Tartu.

References

Crocker, L. and J. Algine, (1986), 'Introduction to classical and modern test theory', *Holt, Rinehart and Windston, Inc.*

Doob, J.L. (1953), 'Stochastic processes', Wiley, New York.

IPCC, (1990), 'Climate change'. The IPCC scientific assessment, Report prepared for IPCC by working group 1; Cambridge University Press.

IPCC, (1992), 'Climate change 1992'. The supplementary report to the IPCC scientific assessment; Cambridge University Press.

Lindley, D.V. (1985), 'Making decisions', Wiley, New York.

Linstone, H.A. and M. Turoff (Eds) (1975), 'The Delphi method', *Techniques and Applications*, Reading, Mass.

Rotmans, J. (1990), 'Image: an integrated model to assess the Greenhouse effect'. Proefschrift Rijksuniversiteit Limburg.

Washington, W.M. and G.A. Meehl, (1989), 'Climate sensitivity due to increased CO_2: Experiments with a coupled atmosphere and ocean general circulation model', *Climate Dyn.* **4**, 1-38.