From Wim van Ackooij, EdF R&D (French Electricity) :

February 24th, 2006

I think that everyone has a more or less clear idea of a model that is not robust. One typically imagines huge models, which need special hardware in order to give results. However a small model is not necessarily a robust one, one might imagine a model where we divide by a small number, which is close to zero due to measurement errors (numerical rounding errors, etc...). A huge model can also be robust and useful. This brings me to believe that a mathematician's contribution should be to make precise what is to be understood with a model being robust.

A three-step approach could be taken in order to do so :

1) Define a new calculus, taking into account uncertainty ;

2) Define what is a model/ robust model ;

3) Make tools to test (2) (probabilistic tools probably).

1) One might imagine a system where numbers are given by two components, the first one being an interval (closed, open, semi-open but never infinite) and the second one being a probability (density) function or measure. We might restrict ourselves to measures that can be written as the integral against a certain density function. In any case this measure should have its supports in the associated interval. We allow Dirac measures for convenience.

[comment, by Bernard Beauzamy : in fact, the probability itself suffices ; its "support" will be the required interval]

This new object translates exactly the ideas of the RMM program, we are not sure of the exact value of something, but we do know something about its distribution around some probable value.

It is easy to see that this is an extension of R (real numbers), since any object with a Dirac measure as second component can be associated with a number of R. Furthermore these new objects are just random variables so we can define multiplication, addition (but not division) etc....

Now any computer program, when calculating a x b can compute (numerically maybe ?) the associated law of this new object a x b. As such the result of the program will automatically contain the uncertainty information required !

2) A mathematical model is just a function F associating to some tuple $(x_1,...,x_n)$ a value $F(x_1,...,x_n)$. Now in order to define what is a robust model I think it is needed to define what is a (Question, Reply) couple, relative to a model. I think we might define this as a given data set $(x_1,...,x_n)$, a set of states $S_1,...,S_m$ and we associate each State S_i with an (unique??) element y_i of R. This can be read as follows, if the model gives us a

certain number as output, we interpret it as being in the State S_i, and therefore we draw the conclusion associated with this State.

Now we can define a model as not being robust if there exists a couple (Q,R), for which we have y_i in $F(x_1,...,x_n)$ and y_j in $F(x_1,...,x_n)$, where $i \neq j$. (Remember that numbers have intervals). This could be interpreted as the model giving an ambiguous result, we may choose the conclusion we want.

[Comments, by Bernard Beauzamy :

Thanks to Wim for this very interesting question : how can we see that a model is robust ? A model is an attempt to reproduce reality by means of formulas. Roughly speaking, it will be considered as robust if the result does not depend too much on the uncertainties on the data, nor on the uncertainties of the laws.

In the simple case considered by Wim, that of a function $F(x_1,...,x_n)$, it means simply that the function F is continuous, as a function of n variables, with a known modulus of continuity. But the fact that the function is simple does not imply that the model is robust : this approximation may be weak. So my impression is that robustness is not a mathematical concept : it describes the way the model adjusts to reality.

Moreover, a model does not need to be a simple function. It may involve time, either discrete or continuous, and it may involve conditional branching, with "if", "then", "else". For instance, the way EdF takes meteorological information into account, in order to manage the production, is by means of complicated "models", which involve complicated "conditional scenarios".]

[Final comments by Wim van Ackooij : In (2) we may include some size parameters, for example a probability level alpha and size parameter epsilon. Instead of demanding the required about the function F, we may ask the following :

"Given α and ε, we can define a model as not being robust if there exists a couple (Q,R), for which we have y_i in $F(x_1,...,x_n)$ and y_j in $F(x_1,...,x_n)$, where $i \neq j$ and the interval of size ε around the mean of x_k contains mass at least $1-\alpha$ for each k."

I think that "if-then-else" statements can be translated into mathematical function by the use of characteristic functions. Loops becomes limits or sums. Mathematical computer programs are often a succession of simple operations such as multiplication or addition.]