More on Models: A sequel to *On the Benefits and Hazards of Models*

By Thomas R. Cuba, Ph.D. August 8th, 2014

Consider the bowling ball. Dropped from the top of the Empire State building it will impact the street below at a very predictable time and with a very predictable velocity. We know because our mathematical model allows us to calculate these outcomes in a very reliable manner. If we were to build a tube from the street to the top of the building, seal it, create a vacuum inside and put in an electronic trigger to release the bowling ball we can eliminate even more uncertainty.

But perhaps I am getting ahead of myself.

In our experiment, we know the weight of the ball and the height of the building. Our model takes these into account. They are the variables, the X and Y, in the model. The force of gravity is also in the model but that comes later. The model does not account for other factors or variables that can affect the outcome: the time and speed that the ball has when it hits the sidewalk. For example, humidity, wind, and air temperature can change the outcome.

Okay, now we can build our tube and eliminate the wind, temperature and humidity. We also improve our ability to measure the time by including the trigger. In short, by eliminating variables, we can increase the accuracy of the result. Note that the vacuum also eliminates wind resistance. Our model matches the real world with great precision and accuracy.

Now let's substitute a crumpled wad of paper for the bowling ball. We can still predict the outcome with a great deal of accuracy because the eliminated variables are still eliminated. If we also get rid of the tube, what we find is that we must include wind resistance, humidity, temperature, and wind in the model but we cannot accurately measure these. So we guess. We take a lot of measurements and use an average of these in the model. Each variable now has an uncertainty. It's like asking someone how much they had to drink. You might get an answer like "Three beers, more or less." In a mathematical model, the *more or less* means the outcome also has a *more or less*.

One more thing. If the model has more variables, and each variable has its own piece of uncertainty, then the *more or less* associated with each of them is multiplied by or added to the *more or less* of the next variable. In summary, models with lots of variables and lots of *more or less* (uncertainty) associated with the actual value of the variables is less reliable and the outcome is much less accurate.

In the first essay (*On the Benefits and Hazards of Models*), the point was made that in many cases the model outcome did not match the real world outcome. In this essay, the point is that a large part of the reason for the mismatch is in the uncertainty of the data. Considering that the numerous models predicting global warming all have hundreds of variables, each with an awful lot of associated *more or less,* it is quite predictable that predicting changes in global average annual temperatures to within a tenth of a degree is unlikely. And I don't even need a model to do it.

I almost forgot. The force of gravity is also in play. To eliminate the *more or less* associated with it, we need to know our latitude on the planet, the moon phase, and what day it is.