How can observations be used to improve models?

In the end, every simulation model is nothing more than an approximating description of reality. This means that there will always be differences between the model and reality. By combining models with observations (data-assimilation), all kinds of interesting options arise, like using models to interpolate between observations or to better understand the measurements.

Observations and models: a powerful pair

Observations and models are a powerful pair. Simulation models are always approximating descriptions of reality. Before a model can be used, it is therefore essential to check whether it appropriately describes the relevant aspects of reality, both quantitatively and qualitatively. For this, observations are indispensible.

Reversely, observations have their limitations too: they represent only a single point in space and time. Whatever happens between or beyond the observation points is, strictly speaking, unknown. Models are used (implicitly or explicitly) to interpolate between observations or to extrapolate observations into the future, i.e. make predictions. Also, models are used to get information about quantities that cannot be observed directly.

There are several ways to combine observations and models. In mathematics, such techniques are referred to as calibration or, more generally, data-assimilation.

Calibration: by hand or automatically

First of all, models can be calibrated using observations. Calibration aims at finding the value for certain parameters in the model so that it best reproduces the observations. Usually, models are calibrated by experts that understand how differences between the model results and the observations are related to the setting of the parameters. But such manual calibration by experts is very labor intensive. On top of that, properly calibrating more than a few parameters at the same time is hardly possible by hand, even for a very bright expert. And even then, there is no guarantee that the settings that have been determined manually are indeed the best settings. For all these reasons, automatic calibration is preferable, certainly if the calibration involves several parameters simultaneously, and needs to be done more than once. OpenDA contains a range of calibration methods.

Computing the input from the output

A special category of methods to calibrate a model uses so called adjoint or inverse models. Simplistically speaking, these models do the reverse computation with respect to the original (or forward) model: they compute the input of the original model (the parameters of the model) from its supposed output (the observations). Constructing an adjoint model usually requires an extensive mathematical analysis and quite some programming. But under certain conditions, it can also be derived automatically from the code of the original model. And besides, there are also inverse methods that can do without an adjoint model. Methods like these are also being developed for OpenDA.

Designing systems
The technique to determine the value of parameters that match a certain model output can also be used conveniently for designing systems: given the properties that a system is supposed to have, calibration methods or inverse methods can be used to determine the required settings of the system. This is an additional type of applications of OpenDA.

Following reality
Rather than adjusting the values of model parameters in a simulation, it is also possible to adapt the simulation results directly to make them better match the observations. For example, for a ground water model the simulated water levels could be replaced by observed values at locations for which observations are available. The model then proceeds the simulation with the more realistic ground water level and will therefore produce better predictions. But modifying the simulation results cannot be done just like that. If the ground water level is modified in one place without adjusting the levels in surrounding computation points, the model is no longer physically correct. Also, the observations are not perfect: they have errors of their own. Mathematics offers a range of techniques that can be used to modify simulation results in a consistent way, taking the uncertainty of both the observations and the model into account. The Kalman filter is probably the classical example of such a technique. OpenDA contains a number of methods that are based on the Kalman filter but do not require the amount of computation that is incurred by the original filter algorithm.