The Kalman Fitler

The tank example was a 0-dimensional simplification of the problem. But the approach and the solution hold more generally. If we can formulate the model state as a state vector $\mathbf{x_b}$ and the observations as a vector $\mathbf{y}$ (of typically much shorter length), we can formulate the same cost function. It then takes the following form ($\mathbf{x}$ ist the best estimate of the state vector of the system):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x_b}) + (\mathbf{y} - H[\mathbf{x}])^T \mathbf{R}^{-1} (\mathbf{y} - H[\mathbf{x}])$$

Here $\mathbf{B}$ is the error covariance matrix of the model (it is a matrix that has the error variances of each element of $\mathbf{x_b}$ in the diagonal; the off-diagonal elements represent the relations between the variables in the form of covariances), and $\mathbf{R}$ is the corresponding error covariance matrix of the observations. H is an operator that reproduces the observation from the model space (e.g., extracts a grid point or interpolates etc.).

The solution resembles the weighted average formulation (right hand side) on the sheet before. In fact, it can be seen as a generalisation of a weighted average:

$$\mathbf{x} = \mathbf{x_b} + \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - H[\mathbf{x_b}])$$

Often the term is $\mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$ is termed „Kalman gain“ $\mathbf{K}$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$$

$\mathbf{H}$ is the Jacobi matrix (matrix of first partial derivatives) of H. The off-diagonal elements of $\mathbf{B}$ control how the update is spread over model space. The update does not only take place at the observation location, but through the covariances between variables it affects the entire model state in a statistically consistent way. Note that not only the state vector is updated, but the covariance matrix $\mathbf{B}$ is similarly updated.

$$\mathbf{B}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$$

In practice, $\mathbf{B}$ is often too large and cannot be estimated precisely. Thus, methods must be found to estimate $\mathbf{B}$ in a useful way.

Data assimilation is mostly used for a system that develops over time. In meteorology, $\mathbf{x_b}$ is typically a short term forecast made with a numerical weather prediction model, which is then compared with the available observations $\mathbf{y}$ to obtain an „analysis“ $\mathbf{x}$. The analysis $\mathbf{x}$ is then used as starting conditions for the next forecast. This yields a sequence of analysis fields. Processing retroactively all observations in this way (including running the numerical weather prediction model for each step) with a fixed system provides what we call a „reanalysis“.