



Data assimilation



Preliminary remarks
Tony O'Hagan

Framework

- ▶ We have a dynamic simulator $f(\mathbf{x}_t, \mathbf{c}, \mathbf{z}_{t+1})$
 - ▶ Inputs: the current state vector \mathbf{x}_t , parameters \mathbf{c} and current forcing inputs \mathbf{z}_{t+1}
 - ▶ Outputs: the new state vector \mathbf{x}_{t+1}
- ▶ At various time points t
 - ▶ We make an observation y_t
 - ▶ That we model as a noisy function of \mathbf{x}_t
 - ▶ I'll assume $y_t = g(\mathbf{x}_t) + \varepsilon_t$
- ▶ What can we learn from such observations?
 - ▶ In order to address that question, we have to acknowledge one important truth ...



All models are wrong

- ▶ The output of the simulator is *not* \mathbf{x}_{t+1}
 - ▶ Even if all the inputs have their true values
 - ▶ In particular, even if we know the true \mathbf{x}_t
 - ▶ We need to acknowledge model discrepancy
- ▶ The correct formulation is
$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{c}, \mathbf{z}_{t+1}) + \delta(\mathbf{x}_t, \mathbf{z}_{t+1})$$
 - ▶ Where $\delta(\mathbf{x}_t, \mathbf{z}_{t+1})$ is model discrepancy
 - ▶ Note that it doesn't formally depend on \mathbf{c}
- ▶ Model discrepancy is hugely important
 - ▶ Without it, there would be little need for data assimilation
 - ▶ And yet almost all theory of DA ignores it!



What can we learn?

- ▶ We have the following system

$$y_t = g(x_t) + \varepsilon_t$$

$$x_t = f(x_{t-1}, c, z_t) + \delta(x_{t-1}, z_t)$$

- ▶ In which there are a number of uncertain quantities
 - ▶ **Conventional DA**
 - ▶ Ignores uncertainty in c and z_t
 - ▶ Treats model discrepancy as a noise term
 - ▶ Only learns about x_t
 - ▶ **But the observations are also informative about c , z_t and $\delta(\cdot, \cdot)$**
 - ▶ DA should be holistic
-



Why does this matter?

- ▶ Because when at every step we need to correct the state vector we should be learning about model discrepancy
 - ▶ And about c (calibration)
- ▶ So that
 - ▶ DA becomes more efficient, but more importantly
 - ▶ So that we can forecast into the future
 - ▶ Or in different contexts
- ▶ DA needs to stop firefighting and learn fire prevention

