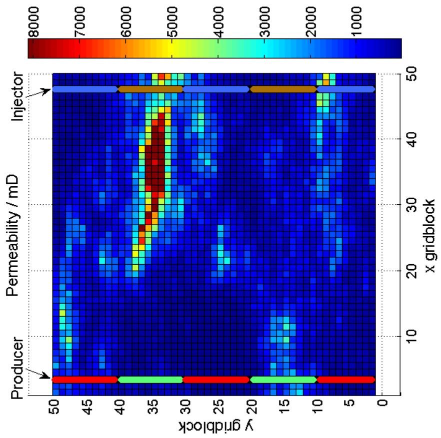
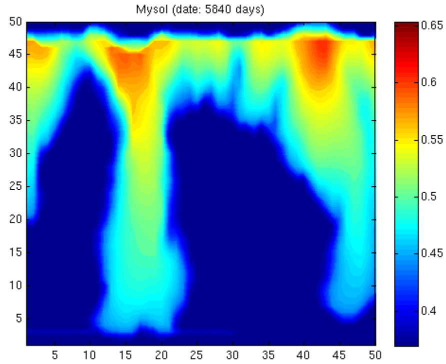
1 Example: Oil Reservoir Optimal Control





2 Problem Setup

Consider a optimal control problem

$$\begin{cases} \dot{x} = F(x, \xi, u) \\ y = g(x) \quad [+\epsilon] \end{cases} \tag{1}$$

with initial condition $x(0) = x_0$. $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $\xi \in \mathbb{R}^n$, $m < n$. ξ is the fixed but unknown. Objective is to maximize

$$J_T = \int_0^T e^{-t\gamma} \phi(x, u) \mathrm{d}t \tag{2}$$

At $t = 0$, the prior on ξ can be modelled by a set of particles $\xi_1, \cdots, \xi^N$. If a myopic approach is adopted, the optimal control will be

$$u^* = \arg\min_u \frac{1}{N} \sum_i J_T \left(x(u, \xi^i; x_0), u \right) , \tag{3}$$

which is equivalent to no observation.

The first approach we thought of is poMDP. We assume $T = \infty$. We discretize time into $t_0 = 0, t_1, \cdots$. If we accounts for the effect of future observations, we use Bellman's equation[1]

$$\begin{aligned} u_t^*(S_t) &= \arg\max_{u_t} \mathbb{E} \left[\left(\phi(x, u_t)\Delta t + e^{-\Delta t \gamma} V(S_{t+1}) \right) \right] \\ V(S_t) &= \mathbb{E} \left[\left(\phi(x, u_t^*)\Delta t + e^{-\Delta t \gamma} V(S_{t+1}) \right) \right] \\ S_{t+1} &= S_{t+1}(S_t, u_t) \end{aligned} \tag{4}$$

S_t is the belief state, i.e. the probability measure of (ξ, x), at t. Notice: not like poMDP, the transition from S_t to S_{t+1} is deterministic for a fixed u_t.

3 A Tentative Alternative Approach?

[1] Warren B. Powell, Approximate Dynamics Programming: Solving the Curses of Dimensionality