Optimization under uncertainty

Purdue ME 597, Distributed Energy Resources

Kevin J. Kircher

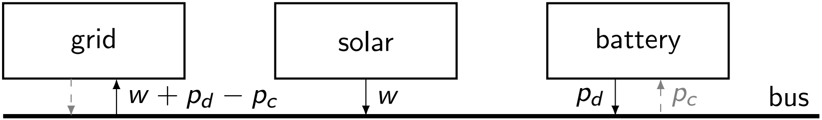
Outline

Uncertainty in optimization

Optimal planning under uncertainty

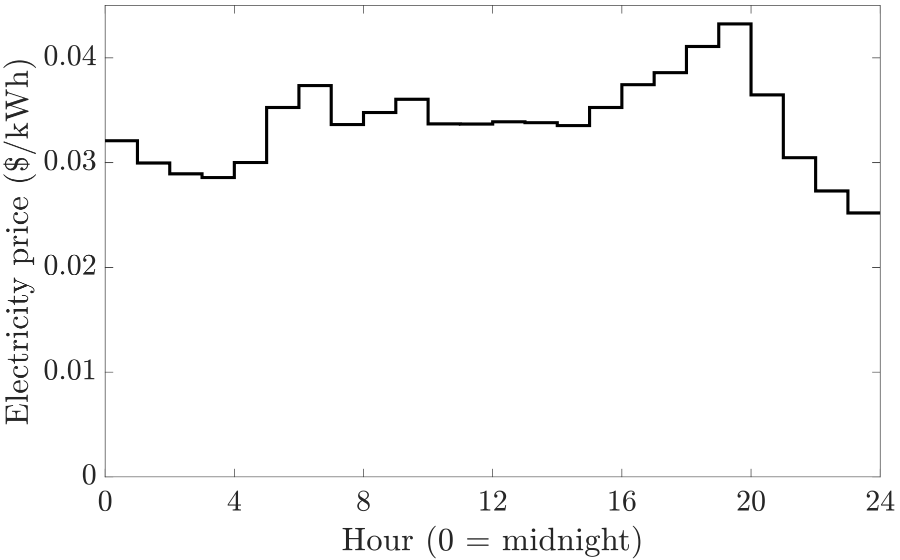
Model predictive control

Example: Energy price arbitrage with grid constraint

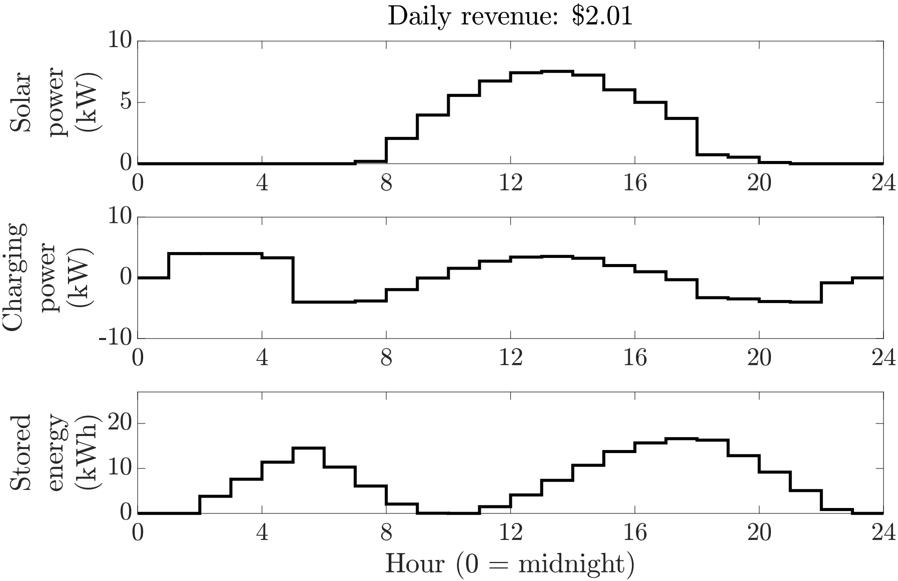


- choose
  - $x = (x(0), \dots, x(K)) \in \mathbf{R}^{K+1}$
  - $p_c = (p_c(0), \dots, p_c(K-1)) \in \mathbf{R}^{K}$
  - $p_d = (p_d(0), \dots, p_d(K-1)) \in \mathbf{R}^{K}$
- to maximize $\Delta t \pi^\top (w + p_d - p_c)$
- subject to $x(0) = x_0$, $x(K) \geq x(0)$, and for $k = 0, \dots, K-1$,
  - $x(k+1) = ax(k) + (1-a)\tau[\eta_c p_c(k) - p_d(k)/\eta_d]$
  - $0 \leq x(k+1) \leq \overline{x}$
  - $0 \leq p_c(k) \leq \overline{p}_c$
  - $0 \leq p_d(k) \leq \overline{p}_d$
  - $|w(k) + p_d(k) - p_c(k)| \leq \overline{p}_g$
- given Δt, π, η_c, η_d, x_0, a, τ, $\overline{x}$, $\overline{p}_c$, $\overline{p}_d$, $\overline{p}_g$, w

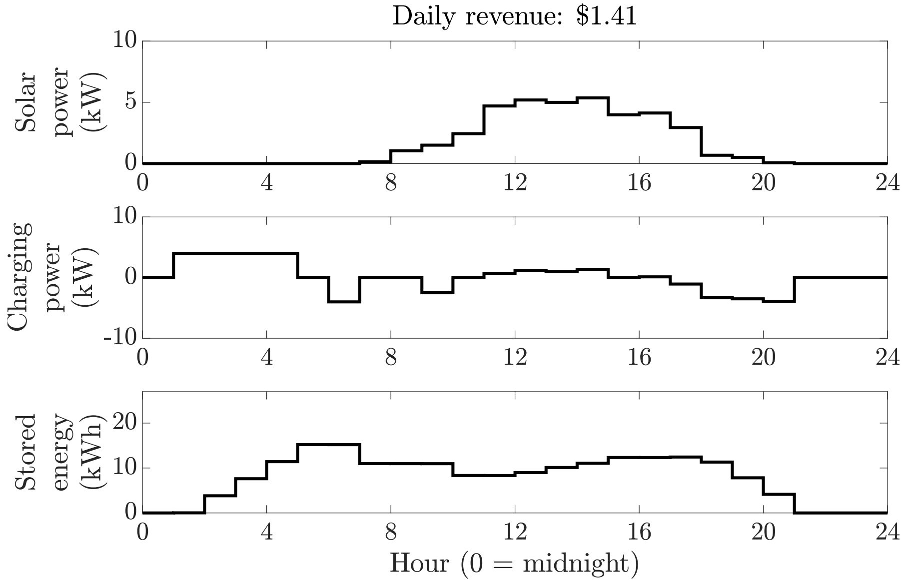
MISO day-ahead electricity price



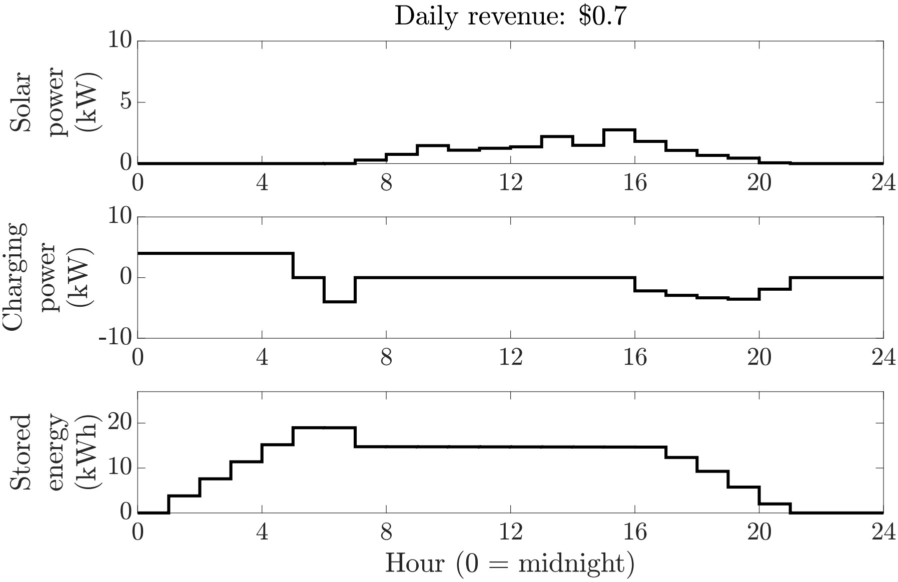
Omniscient optimization for a clear day



Omniscient optimization for a partly cloudy day



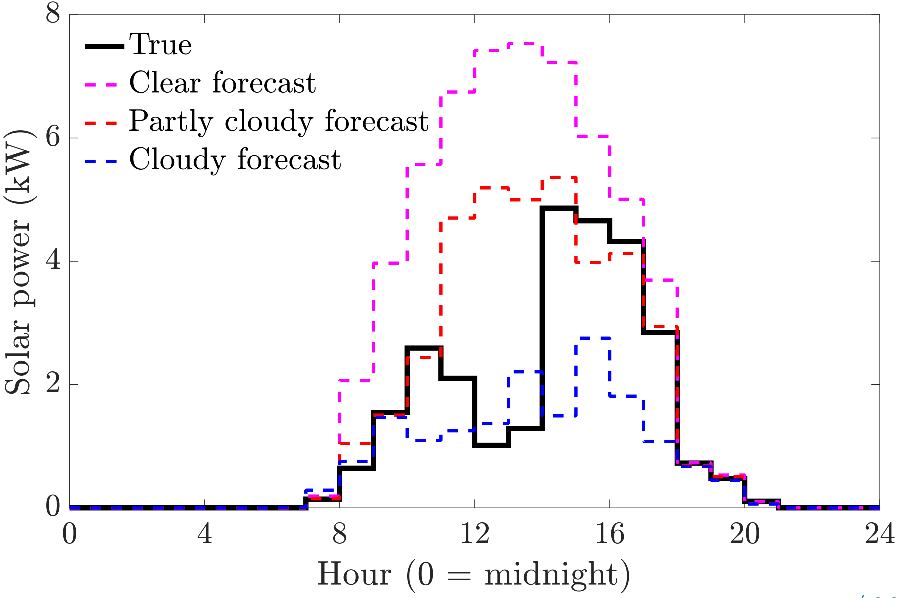
Omniscient optimization for a cloudy day



What if we're not omniscient?

- the omniscient examples above assume
  - perfect knowledge of the model structure and parameters
  - perfect measurement of the initial stored energy
  - perfect foreknowledge of solar power outputs and energy prices
- in reality, some or all of these assumptions may be bad
- for example, suppose
  - our solar forecast is imperfect
  - we use the forecast to make a plan and stick to it all day

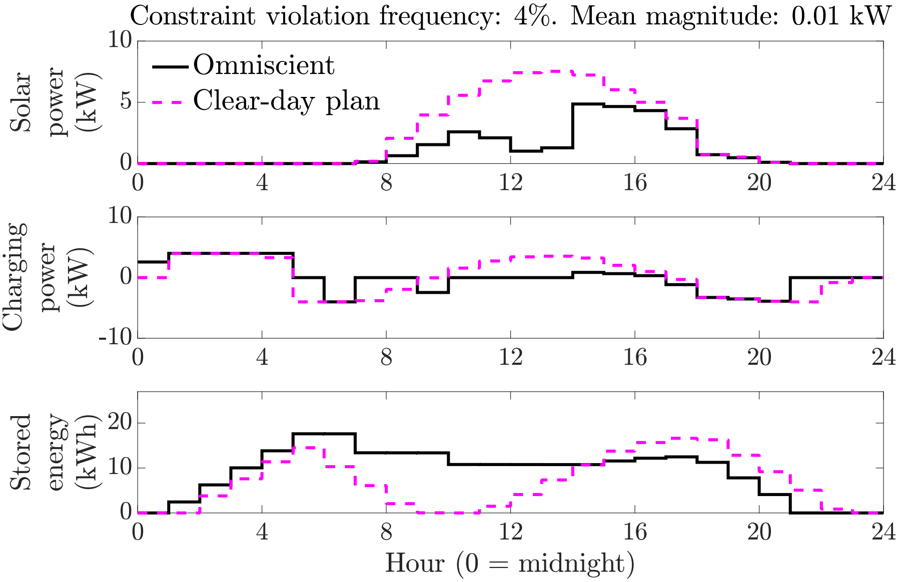
True and forecasted solar powers



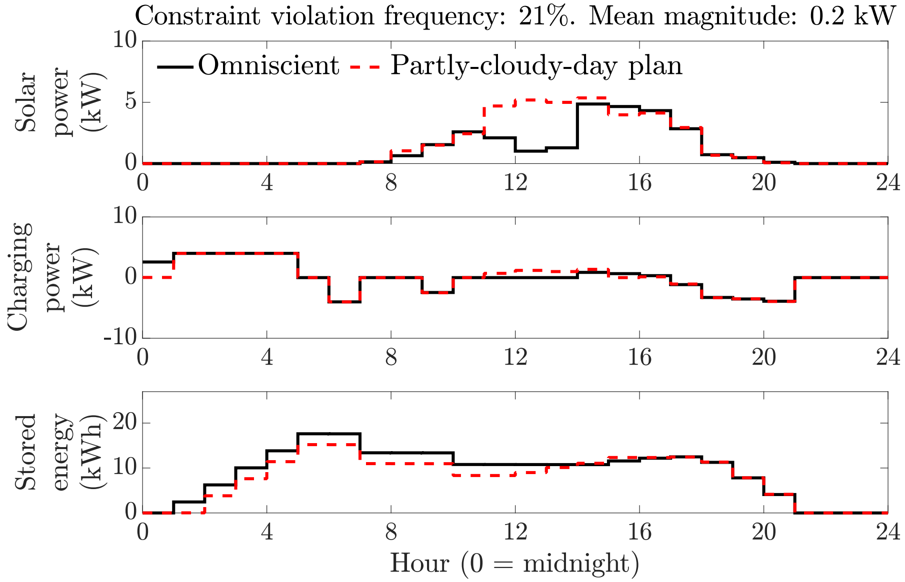
Planning with a forecast $\hat{w}$ of solar power w

- choose
  - $x = (x(0), \dots, x(K)) \in \mathbf{R}^{K+1}$
  - $p_c = (p_c(0), \dots, p_c(K-1)) \in \mathbf{R}^K$
  - $p_d = (p_d(0), \dots, p_d(K-1)) \in \mathbf{R}^K$
- to maximize $\Delta t \pi^\top (\hat{w} + p_d - p_c)$
- subject to $x(0) = x_0$, $x(K) \geq x(0)$, and for $k = 0, \dots, K-1$,
  - $x(k+1) = ax(k) + (1-a)\tau[\eta_c p_c(k) - p_d(k)/\eta_d]$
  - $0 \leq x(k+1) \leq \overline{x}$
  - $0 \leq p_c(k) \leq \overline{p}_c$
  - $0 \leq p_d(k) \leq \overline{p}_d$
  - $|\hat{w}(k) + p_d(k) - p_c(k)| \leq \overline{p}_g$
- given Δt, π, η_c, η_d, x_0, a, τ, $\overline{x}$, $\overline{p}_c$, $\overline{p}_d$, $\overline{p}_g$, $\hat{w}$

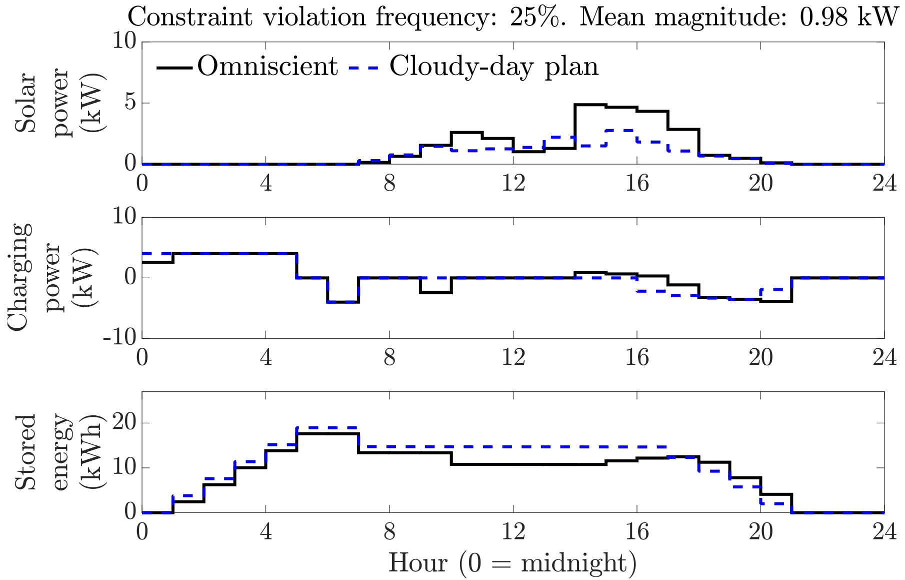
Results from planning on 'clear', $\hat{w} = w^{\text{clr}}$



Results from planning on 'partly cloudy', $\hat{w} = w^{pc}$



Results from planning on ‘cloudy’, $\hat{w} = w^{\text{cld}}$



Effects of forecast errors

forecast	violation frequency	conditional average violation magnitude	revenue
perfect	0%	0 kW	\$1.07
‘clear’	4%	0.01 kW	\$1.00
‘partly cloudy’	21%	0.2 kW	\$1.08
‘cloudy’	25%	0.98 kW	\$1.06

how to hedge against risks of forecast errors?

Outline

Uncertainty in optimization

Optimal planning under uncertainty

Model predictive control

Risk-neutral planning

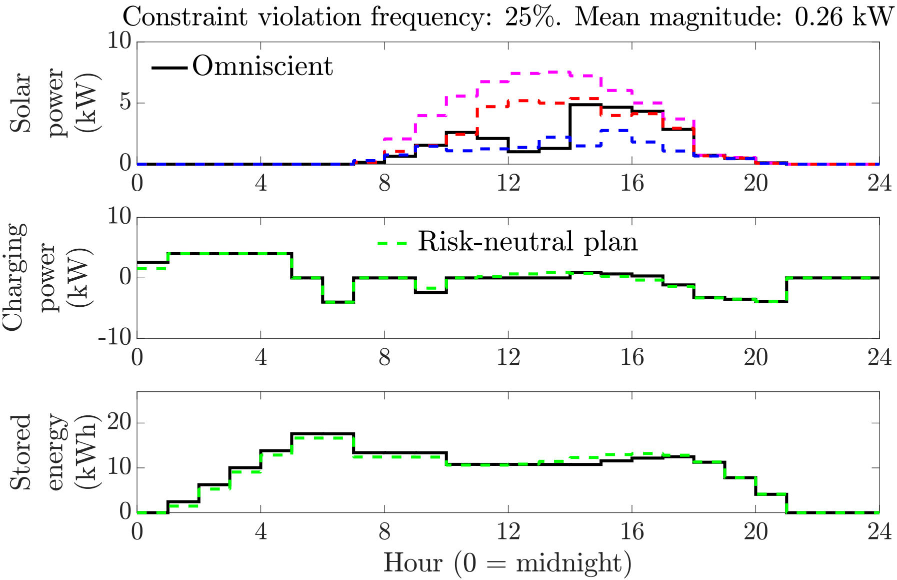
- recall that w only influences optimization through the
  - objective: maximize $\Delta t \pi^\top (w + p_d - p_c)$
  - grid interconnection constraint: $|w(k) + p_d(k) - p_c(k)| \leq \overline{p}_g$
- a <u>risk-neutral</u> planning approach
  - maximizes the expected revenue $\mathbf{E}\left[\Delta t \pi^\top (w + p_d - p_c)\right]$
  - replaces the ambiguous constraint

$$|w(k) + p_d(k) - p_c(k)| - \overline{p}_g \leq 0$$

  by the expected-value constraint

$$\mathbf{E}\left[|w(k) + p_d(k) - p_c(k)| - \overline{p}_g\right] \leq 0$$

Risk-neutral planning with w^{clr}, w^{pc}, w^{cld}, equally likely



Risk-neutral optimization more generally

- consider the (ambiguous) problem with uncertain vector δ
  - choose $x \in \mathbf{R}^n$
  - to minimize $f_0(x, \delta)$
  - subject to $f_1(x, \delta) \leq 0, \ldots, f_m(x, \delta) \leq 0$
  - given $f_0, \ldots, f_m$ and... something about δ ?!
- a risk-neutral approach models δ as random and solves
  - choose $x \in \mathbf{R}^n$
  - to minimize $\mathbf{E}_\delta\, f_0(x, \delta)$
  - subject to $\mathbf{E}_\delta\, f_1(x, \delta) \leq 0, \ldots, \mathbf{E}_\delta\, f_m(x, \delta) \leq 0$,
  - given $f_0, \ldots, f_m$ and the distribution of δ
    (and the ability to compute the expectation integrals)

(Maximally) risk-averse planning

a (maximally) risk-averse approach

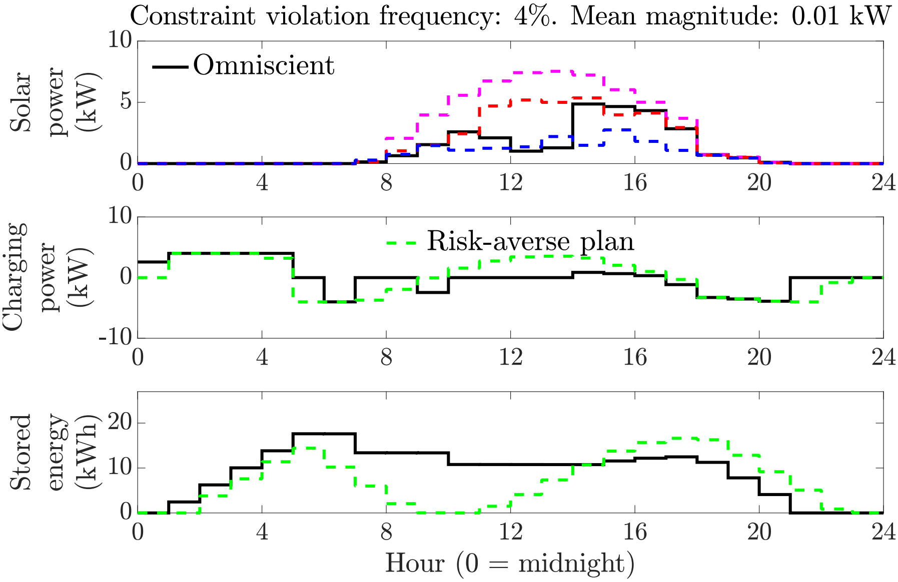
- maximizes the worst-case revenue $\min_w \left[\Delta t \pi^\top (w + p_d - p_c)\right]$
- replaces the ambiguous constraint

$$|w(k) + p_d(k) - p_c(k)| - \overline{p}_g \leq 0$$

by the worst-case constraint

$$\max_{w(k)} \left[|w(k) + p_d(k) - p_c(k)| - \overline{p}_g\right] \leq 0$$

Risk-averse planning with only w^{clr}, w^{pc}, w^{cld} possible



(Maximally) risk-averse optimization more generally

- consider the (ambiguous) problem with uncertain vector δ
  - choose $x \in \mathbf{R}^n$
  - to minimize $f_0(x, \delta)$
  - subject to $f_1(x, \delta) \leq 0, \ldots, f_m(x, \delta) \leq 0$
  - given $f_0, \ldots, f_m$ and... something about δ ?!
- a maximally risk-averse approach solves
  - choose $x \in \mathbf{R}^n$
  - to minimize $\max_\delta f_0(x, \delta)$
  - subject to $\max_\delta f_1(x, \delta) \leq 0, \ldots, \max_\delta f_m(x, \delta) \leq 0$
  - given $f_0, \ldots, f_m$ and the support of δ
    (and the ability to compute the maxima over δ)

Planning with mixed risk measures

it sometimes makes sense to mix risk measures; for example:

- risk-neutral objective: maximize $\mathbf{E}\left[\Delta t \pi^\top (w + p_d - p_c)\right]$
- (maximally) risk-averse constraints:

$$\max_{w(k)} \left[|w(k) + p_d(k) - p_c(k)| - \overline{p}_g\right] \leq 0$$

- in solar/battery example, results resemble risk-averse approach

Other risk measures

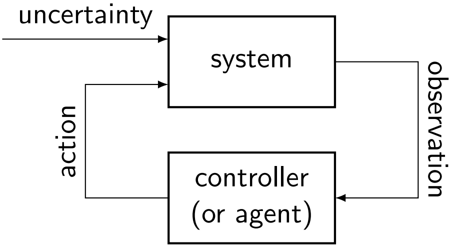
- a risk measure $\mathcal{R}$ disambiguates ambiguous constraints

$$f(x,\delta) \leq 0 \text{ (ambiguous)} \longrightarrow \mathcal{R}[f(x,\delta)] \leq 0 \text{ (unambiguous)}$$

- the expected value, $\mathcal{R} = \mathbf{E}_\delta$, is risk-neutral
- the worst-case value, $\mathcal{R} = \max_\delta$, is maximally risk-averse
- other risk measures interpolate between these extremes
  - value-at-risk
  - conditional value-at-risk
  - entropic value-at-risk
  - others...

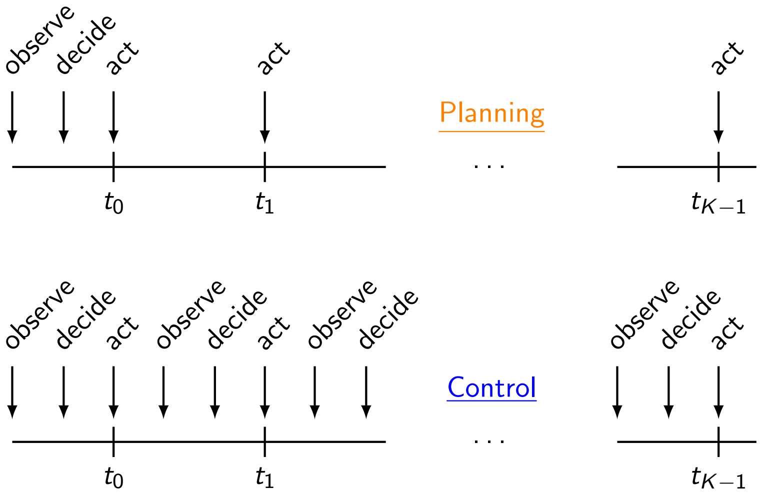
What if we don't have to rigidly stick to our plan?

- often, we can update actions based on new information
- this option is called **recourse**
- exercising it gives rise to **feedback**



- many names for sequential decision-making under uncertainty
  - {feedback, stochastic, optimal, robust} control
  - multi-stage stochastic programming (with recourse)
  - Markov decision processes, reinforcement learning

Planning vs. control



Outline

Uncertainty in optimization

Optimal planning under uncertainty

Model predictive control

The general MPC algorithm

repeat:

- **observe** the system
- **predict** uncertain influences on the system over a receding horizon
- **decide** a plan of action under those predictions
- **act** on the plan for a while

The MPC planning problem at time k

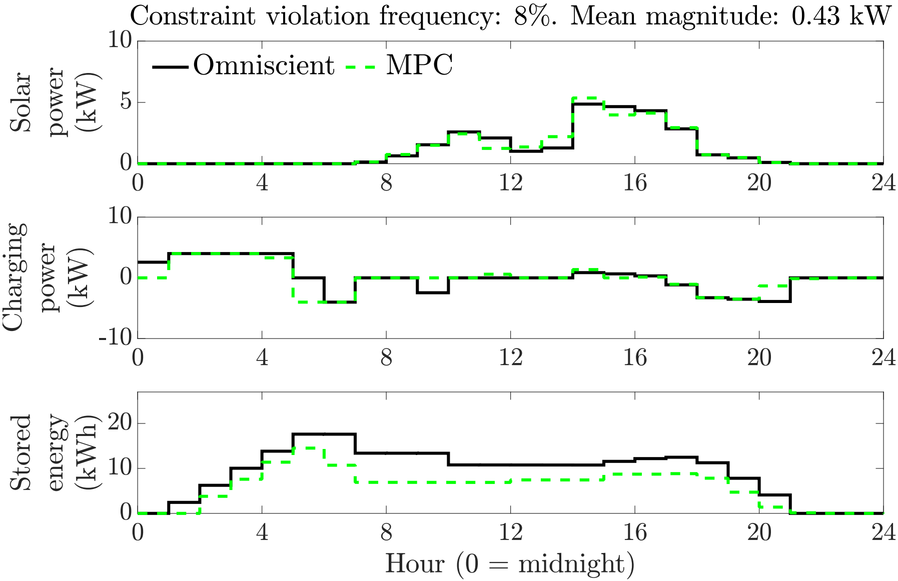
- pick a planning horizon H
- make a prediction $\hat{w}(h|k)$ of $w(k+h)$ for $h = 0, \ldots, H-1$
- choose
  - $(x(0|k), \ldots, x(H|k)) \in \mathbf{R}^{H+1}$
  - $(p_c(0|k), \ldots, p_c(H-1|k)) \in \mathbf{R}^{H}$
  - $(p_d(0|k), \ldots, p_d(H-1|k)) \in \mathbf{R}^{H}$
- to maximize $\Delta t \sum_{h=0}^{H-1} \pi(k+h)(\hat{w}(h|k) + p_d(h|k) - p_c(h|k))$
- subject to $x(0|k) = x(k)$ and for $h = 0, \ldots, H-1$,
  - $x(h+1|k) = ax(h|k) + (1-a)\tau[\eta_c p_c(h|k) - p_d(h|k)/\eta_d]$
  - $0 \leq x(h+1|k) \leq \overline{x}$
  - $0 \leq p_c(h|k) \leq \overline{p}_c$
  - $0 \leq p_d(h|k) \leq \overline{p}_d$
  - $|\hat{w}(h|k) + p_d(h|k) - p_c(h|k)| \leq \overline{p}_g$

  (& possibly a terminal cost or constraint, e.g. $x(H|k) \geq x(k)$)
- given Δt, π, η_c, η_d, x_0, a, τ, $\overline{x}$, $\overline{p}_c$, $\overline{p}_d$, $\overline{p}_g$, $\hat{w}(0|k), \ldots, \hat{w}(H-1|k)$

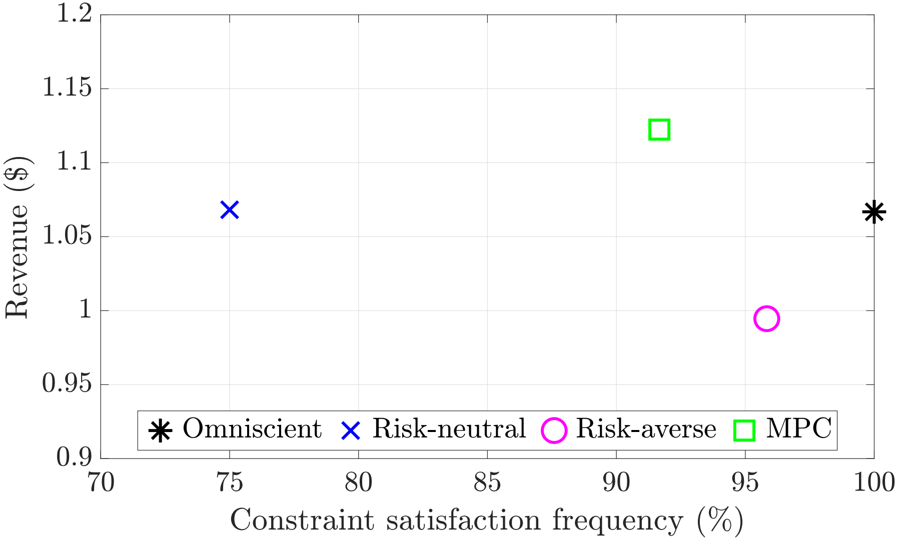
A simple prediction algorithm

- at each time k,
  - recall the last solar observation, $w(k-1)$
  - compare it to $w^{\mathrm{clr}}(k-1)$, $w^{\mathrm{pc}}(k-1)$, and $w^{\mathrm{cld}}(k-1)$
  - denote the closest match by $w^\star(k-1)$
  - predict $\hat{w}(h|k) = w^\star(k+h)$ for $h = 0, \ldots, H-1$
- in words: predict that the current conditions will persist
- you can imagine much more sophisticated predictors than this

MPC results


Constraint violation frequency: 8%. Mean magnitude: 0.43 kW
Omniscient
MPC
Solar power (kW)
Charging power (kW)
Stored energy (kWh)
Hour (0 = midnight)
0
4
8
12
16
20
24
10
5
-10
20

Balancing performance and robustness



Notes on MPC

- although it uses optimization, **MPC is generally not optimal**
  (other algorithms can, and sometimes do, perform better)
- but it often performs unreasonably well in practice
- models, forecasts, and state estimates don't have to be great
- MPC dates back to the 1980s and many fields use it
  (chemical processing, robotics, finance, supply chain management, ...)
- there are many MPC variants
  (nonlinear, stochastic, robust, distributed, adaptive, ...)
- MPC is overkill for many problems