NREL
Transforming ENERGY

Adaptive computing and multi-fidelity strategies for control, design and scale-up of renewable energy applications

Marc Day and Deepthi Vaidhynathan
Computational Science Center
NREL

Photo by Dennis Schroeder, NREL 55200

Risk in the scale-up of complex systems

Scale-up: Extending systems and processes that were developed in the laboratory to function in the real world

Device and process scale-up comes with significant technical challenges and risk

Typical challenges:

- Data-driven models perform best when interpolating, **extrapolation is inherently uncertain,** and therefore **risky**
- Increasing ranges of scale (spatial, temporal) often lead to new/enriched physics
- High-fidelity physics-based models may capture new physics, but are typically too expensive for design/optimization work
- Operational regimes of existing experiments are limited, and new experiments are expensive

Image Credit: Dennis Schroeder, NREL 63958

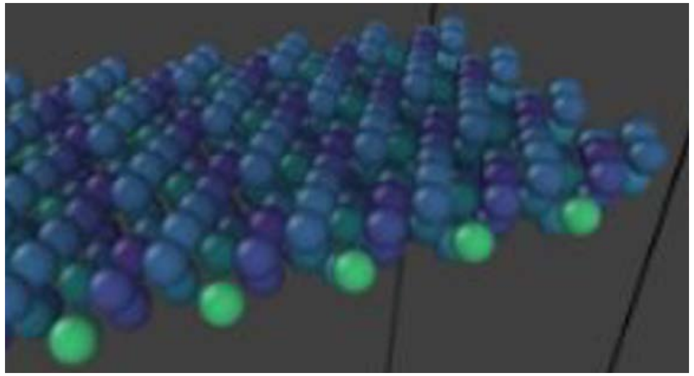
Adaptive computing: A holistic modeling approach to address scale-up challenges across NREL application domains

NREL LDRD FY23-25

PI: M. Day

Adaptive Computing

Orchestration of a multi-fidelity model hierarchy and/or experiment campaign to arrive at the best goal-based solution with well-characterized uncertainty given finite resources



Key algorithmic components of Adaptive Computing

Adaptive Computing

Orchestration of a multi-fidelity model hierarchy and/or experiment campaign to arrive at the best goal-based solution with well-characterized uncertainty given finite resources

Optimize the use of finite resources to to achieve a specific science goal

Goal-oriented solutions

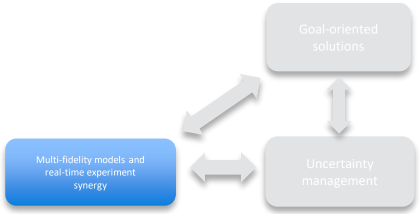
Multi-fidelity models and real-time experiment synergy

Connect models with experiments to drive experiment design and data acquisition needs

Uncertainty management

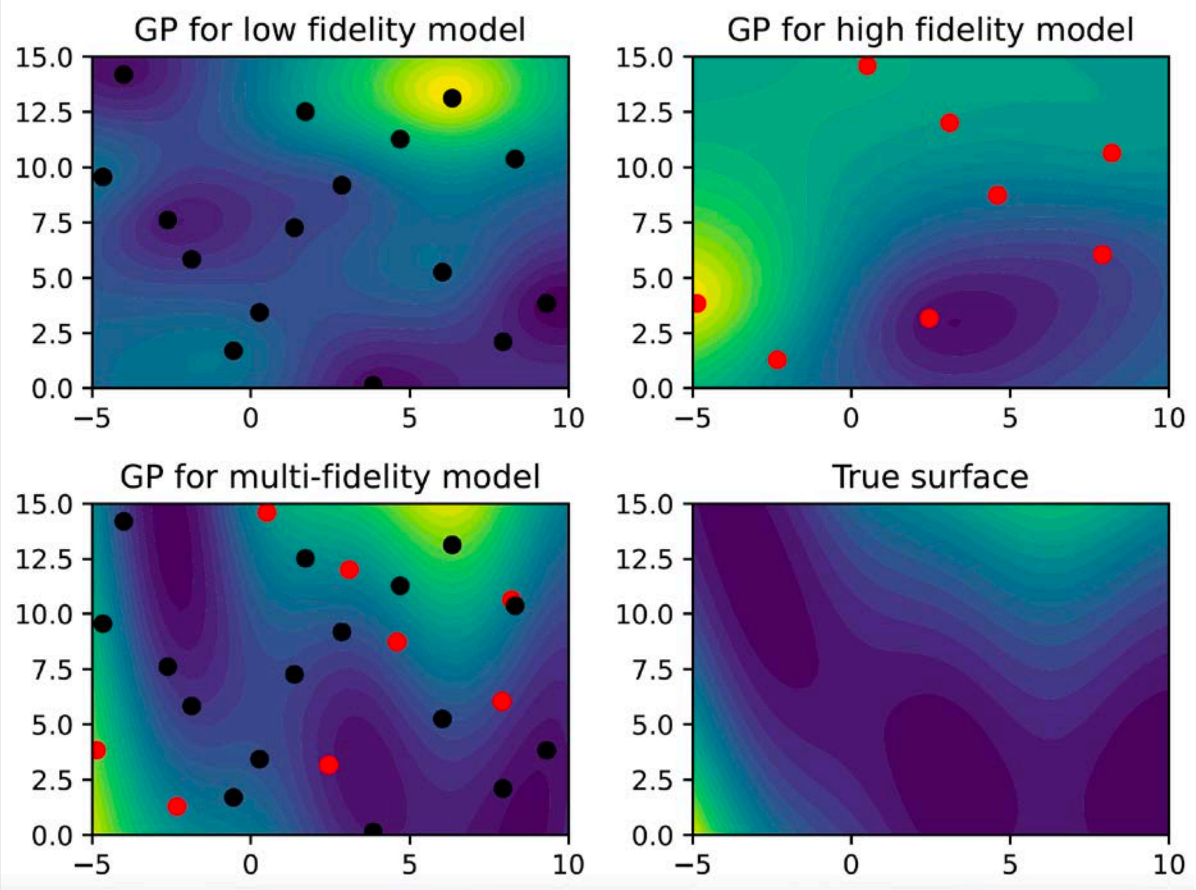
Control of extrapolation uncertainty through targeted active learning

Key Capability: Multi-fidelity modeling



Most applications feature an assortment of models of widely varying fidelities, developed for different purposes:

- **Experiment**: “Truth”, but limited operational regime
- **High-fidelity:** Physics-based (PDE/ODE models), can be extremely costly
- **Low fidelity**: reduced physics, coarser meshes, cheaper
- **Data-driven surrogates:** AI/ML, PINNs, Gaussian Processes (GPs). Typically, very cheap



Exploiting information from multiple fidelity levels can increase surrogate accuracy

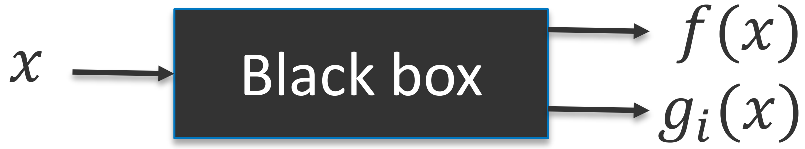
Black-box expensive optimization: high fidelity

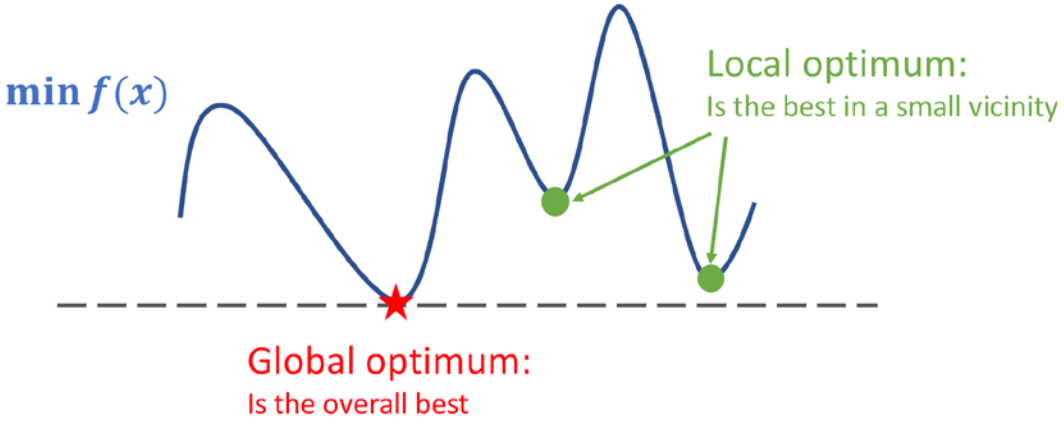
$$\min f(x)$$
$$s.t.\ g_i(x) \leq 0, i = 1, \ldots, I$$
$$x \in \Omega$$

Objective function to minimize

Constraint functions

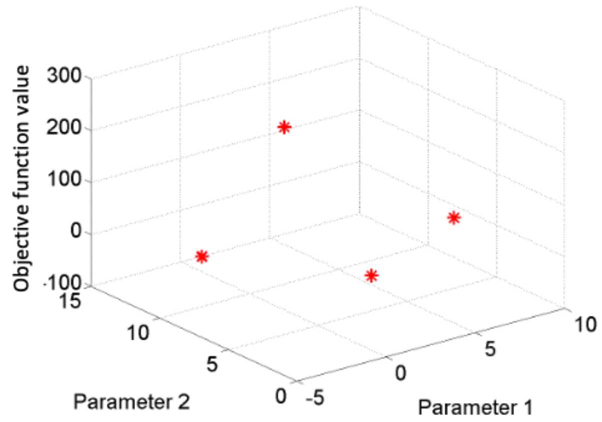
Parameter domain



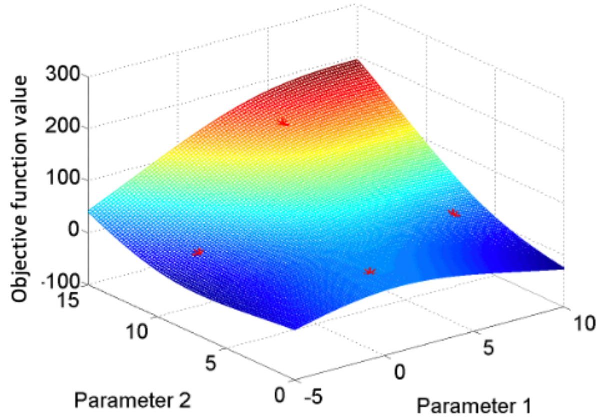


$$\cancel{\frac{df}{dx}}$$

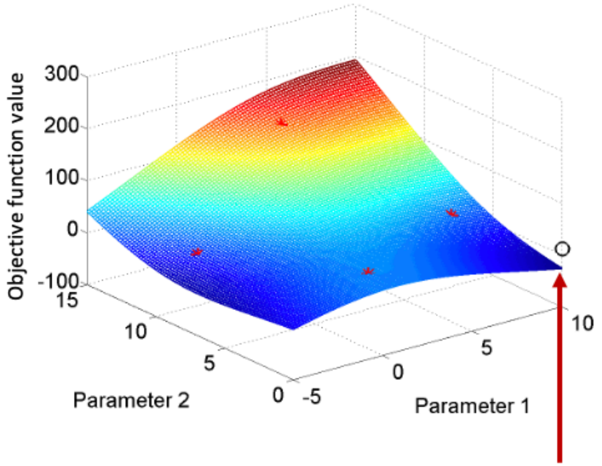
Surrogate model guided sampling



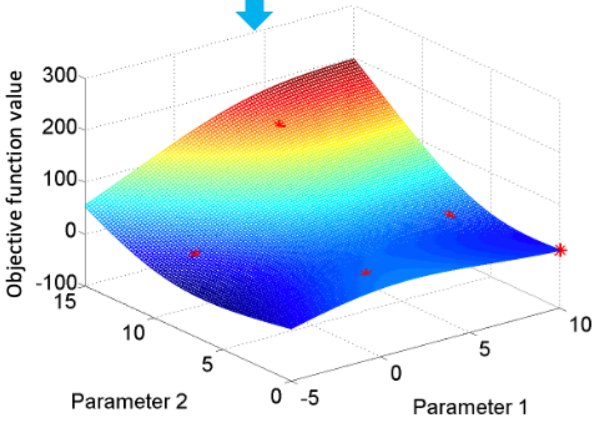
Initial experimental design



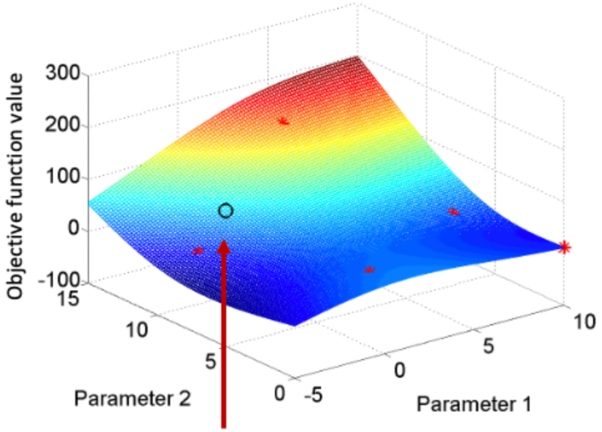
Compute surrogate surface



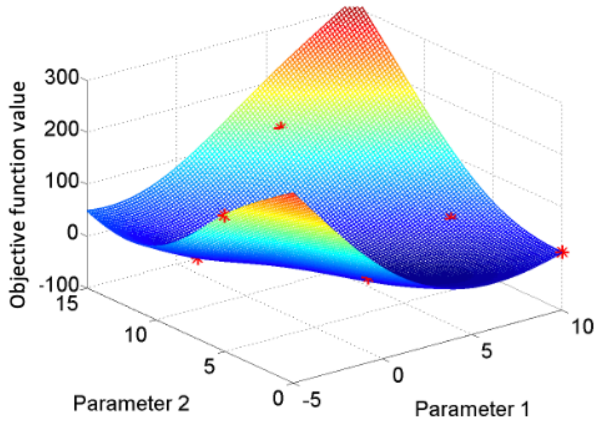
Select and evaluate new point



Update surrogate surface



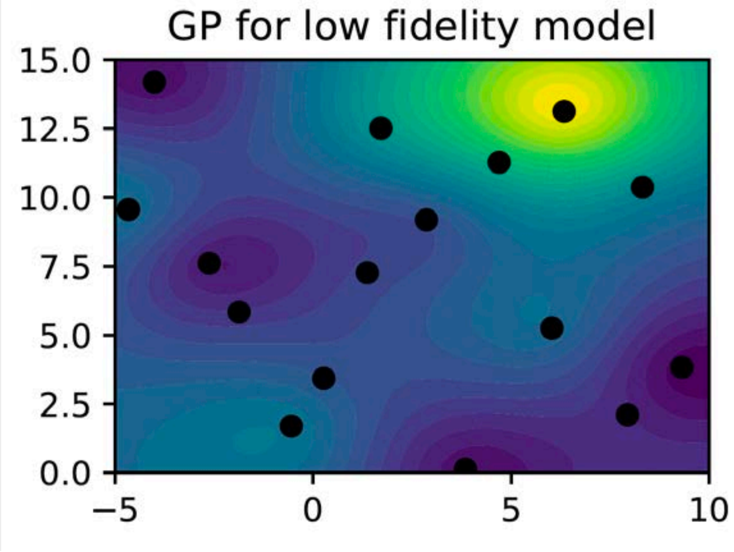
Select and evaluate new point

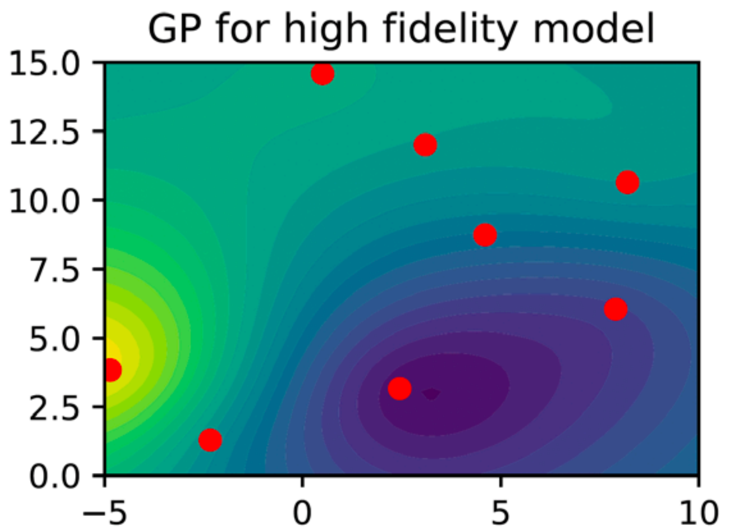


Update surrogate surface

...

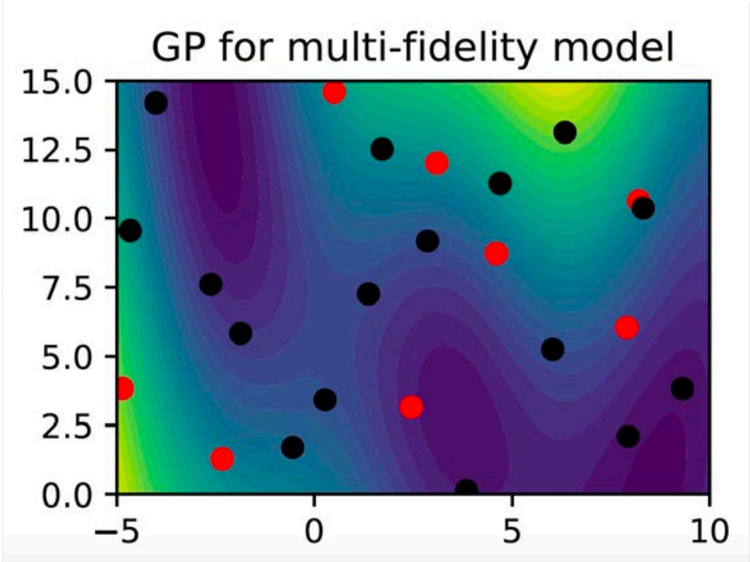
Combining information from multiple fidelities into one model

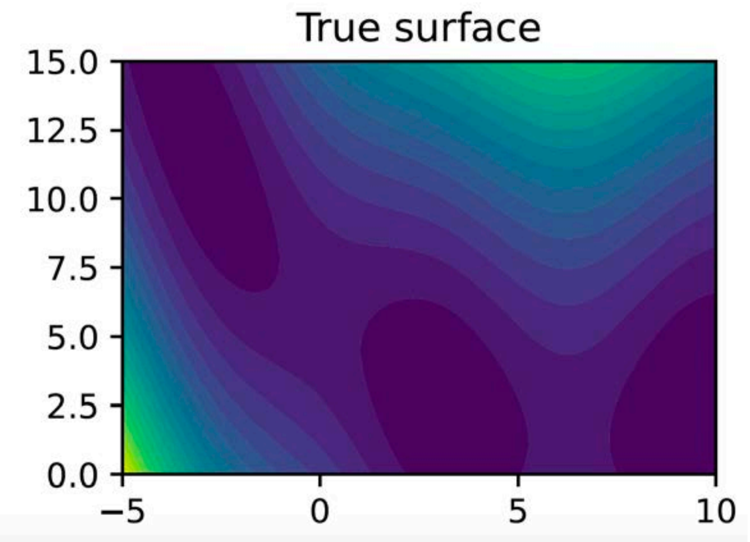




Red = high-fidelity evaluations
Black = Lower fidelity evaluations

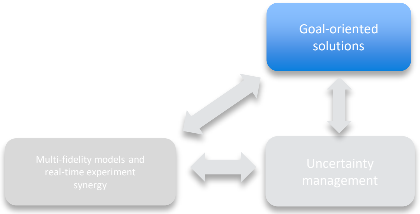
Python package surrogate modeling toolbox (SMT)





Combining high and lower fidelity information can lead to better approximation surface (compare to true contours) for similar cost

Goal-oriented sampling methods



Expected improvement balances local and global search

$$\mathbb{E}(I)(\mathbf{x}) = s(\mathbf{x})\left(\nu\Phi(\nu) + \phi(\nu)\right), \quad \nu = \frac{f^{\text{best}} - m_{\text{GP}}(\mathbf{x})}{s(\mathbf{x})}$$

Probability of improvement – mostly local search

$$\mathbb{P}(I)(\mathbf{x}) = \Phi\left(\frac{f^{\text{best}} - m_{\text{GP}}(\mathbf{x})}{s(\mathbf{x})}\right)$$

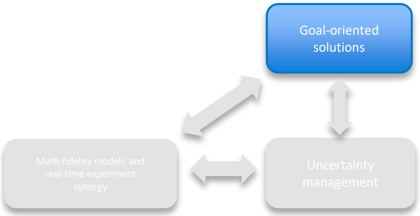
Lower confidence bound

$$LCB(\mathbf{x}) = m_{\text{GP}}(\mathbf{x}) - \kappa s(\mathbf{x})$$

m_{GP} is the prediction from the surrogate (e.g., GP)
s is the standard deviation of the GP predictions
f^{best} is the best function value found so far
ϕ, Φ are the normal density and cumulative distribution

κ – adjustable parameter

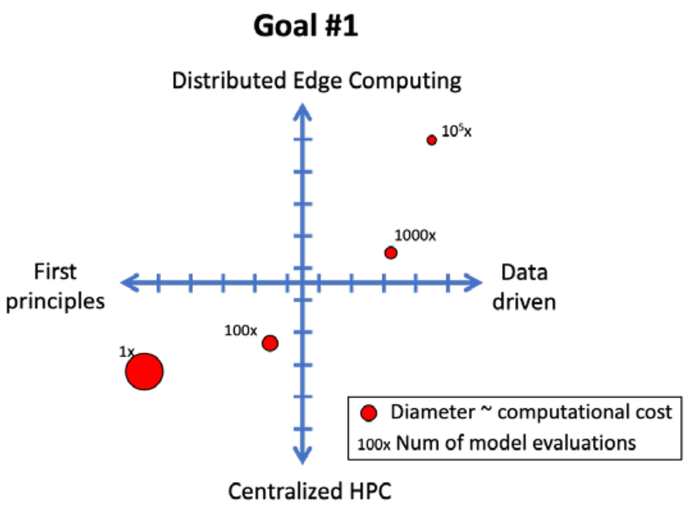
Leveraging distributed heterogenous data sources

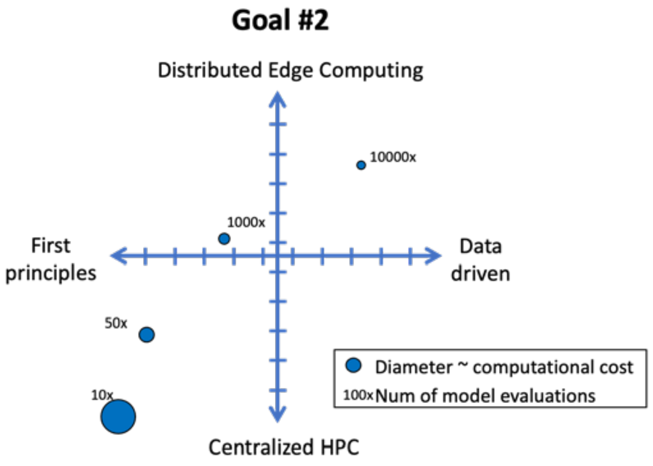


Investigations may target a variety of design goals, e.g.
- Maximize energy efficiency
- Minimize down times
- Minimize operating costs
- Minimize pollutants
- …

Each of these may lead to different requirements
- Accuracy
- Uncertainty
- Reproducibility
- …

…and thus a different campaign of data acquisition, including a mix of:
- Experimental data gathering
- Surrogate/ROM evaluations (including any required training)
- High-fidelity simulations





Leveraging diverse compute resources

Experiment

Edge

Cloud

Datacenter

Resource Management

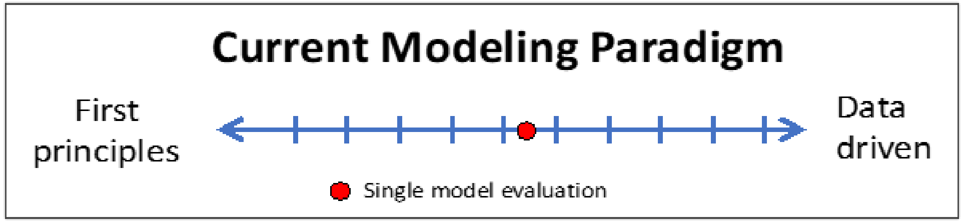
Resource manager:
how much and what kind of resources do I have available?

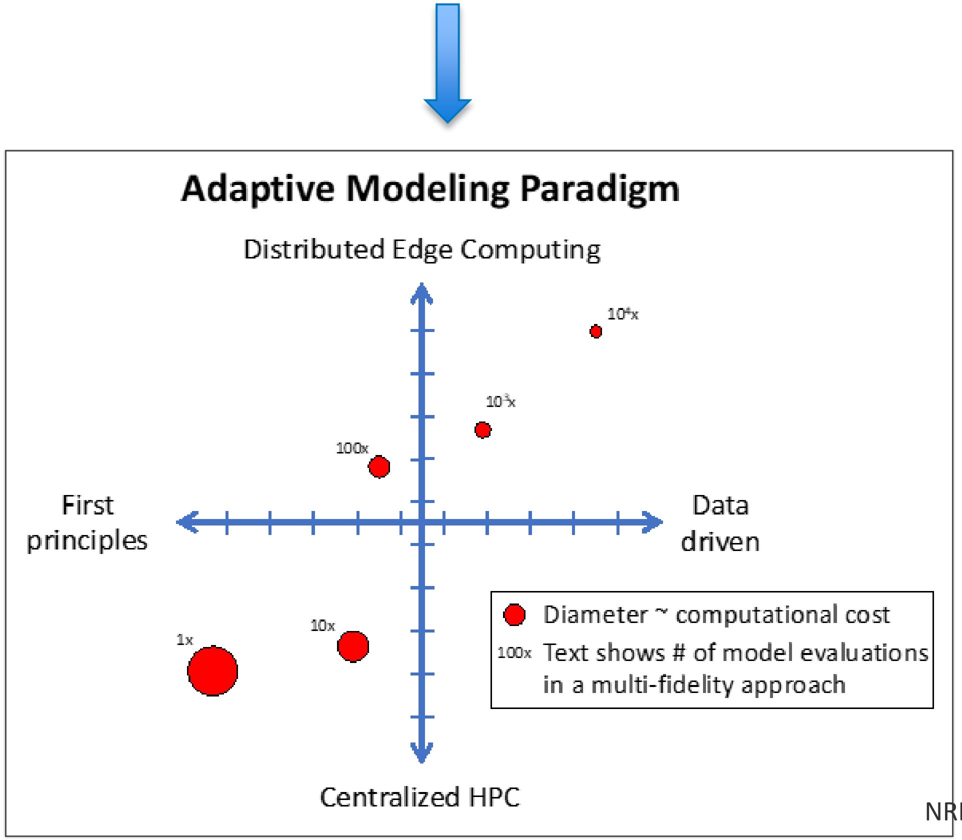
Solve stochastic discrete optimization problem

Optimal computing strategy driven by specific output quantity of interest

Compute resource optimization problem

- What resources are available when?
- Formulate as optimization problems with stochasticity
- Implement solutions as constraints for multi-fidelity sampling
- Eventually must exploit asynchronous parallel computations
- Enumerate the user-defined simulation types (fidelity levels)
- Possible hardware configurations (# of CPUs, GPUs)
- Corresponding calculation duration
- Measurement noise estimate (aleatoric uncertainty)





Compute resource allocation with stochasticity

1. Select sample points (e.g., maximize EI with multi-start; candidate point approach)

$$\{x_1, x_2, \ldots, x_N\} \in \underset{x \in \mathcal{X}}{argmax}\, a(x)$$

2. Get total resource limit T and per level resource limit T_j and allocate compute resources

$$\max_{\mathrm{b_{ji}} \in \{0,1\}^{\mathrm{k}}} \sum_{j=1}^{J} r_j(x_i) \ * b_{ji}$$

r_j the benefit of evaluating x_i at fidelity level j, e.g.,
r_j captures accuracy or other QoI

$$b_{ji} = \begin{cases} 1 \text{if } x_i \text{ evaluated with fidelity level } j \\ 0 \text{ else} \end{cases}$$

$J * N$ binaries to optimize

$$\sum_{j=1}^{J} \sum_{i=1}^{N} b_{ji} * t_j(x_i, \zeta_{ji}) \leq \ T$$

Total resource restriction

t_j resource consumption at level j

$\zeta_{ij} \sim ?$

$$\sum_{i=1}^{N} b_{ji} * t_j\left(x_i, \zeta_{ji}\right) \leq \ T_j \quad \forall j$$

Resource restriction on fidelity level j

Uncertainty management and extrapolation

Goal-oriented solutions

Multi-fidelity models and real-time experiment synergy

Uncertainty management

Leverage ML models across system hierarchy

- Small data: Gaussian processes
- Large data: Bayesian NNs
- Incorporating data across fidelities

guide

Adaptive learning and data acquisition for QoI and model uncertainty

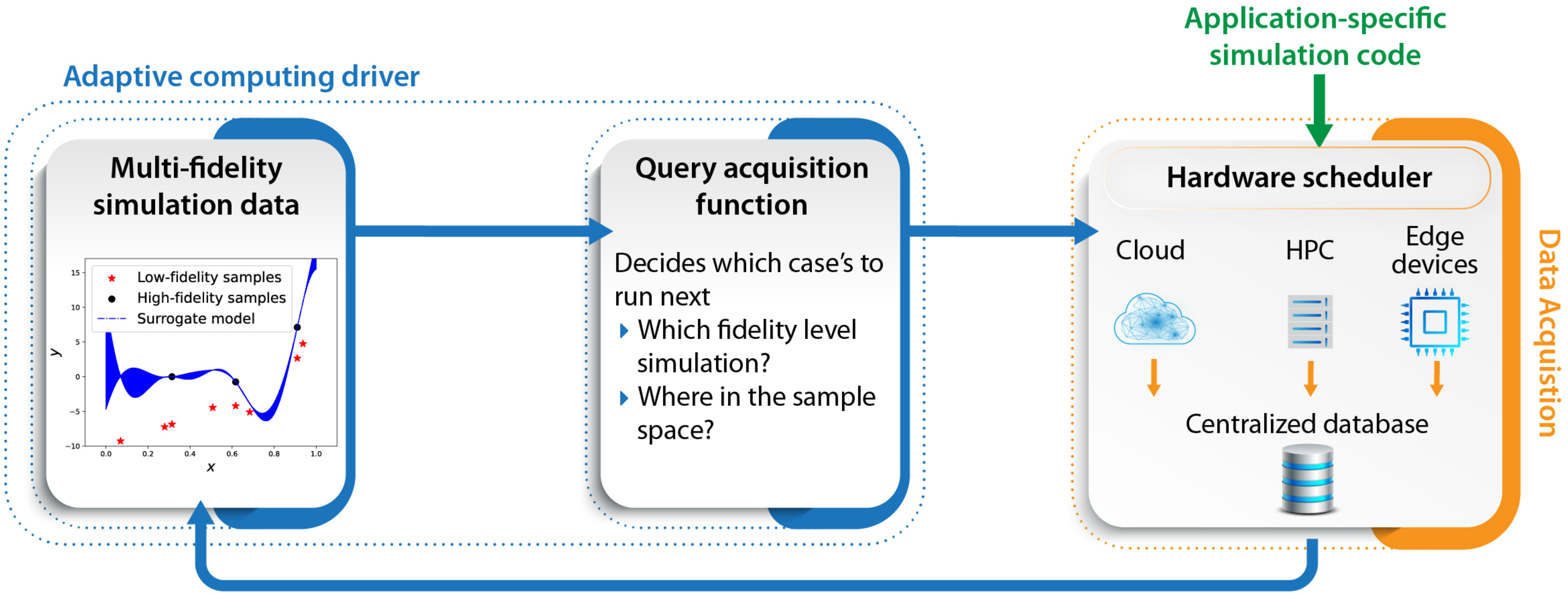
update

send queries

Sample evaluation: Experiments/high-fidelity evaluations and surrogates

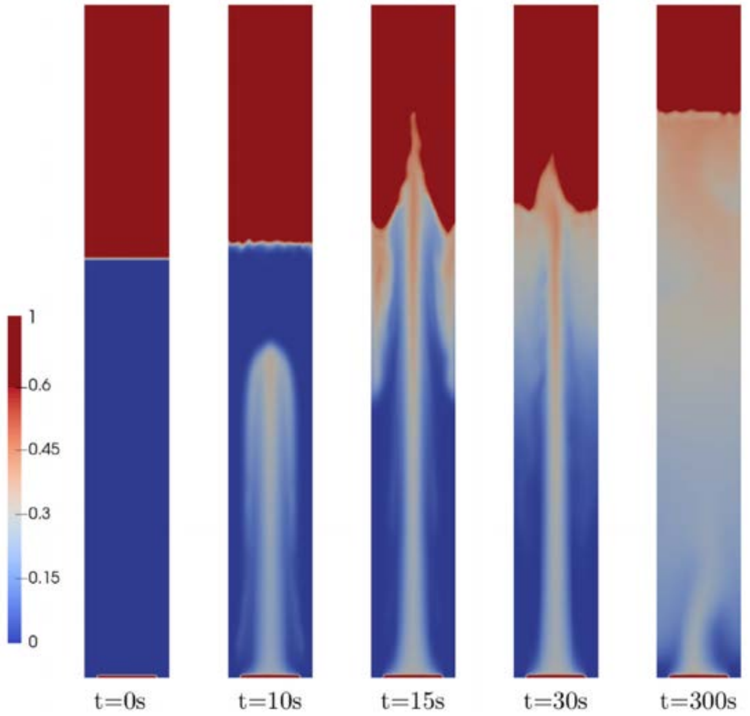
Requires implementation of a library of ML models, model validation techniques and decision frameworks for model updates

Adaptive Computing Software



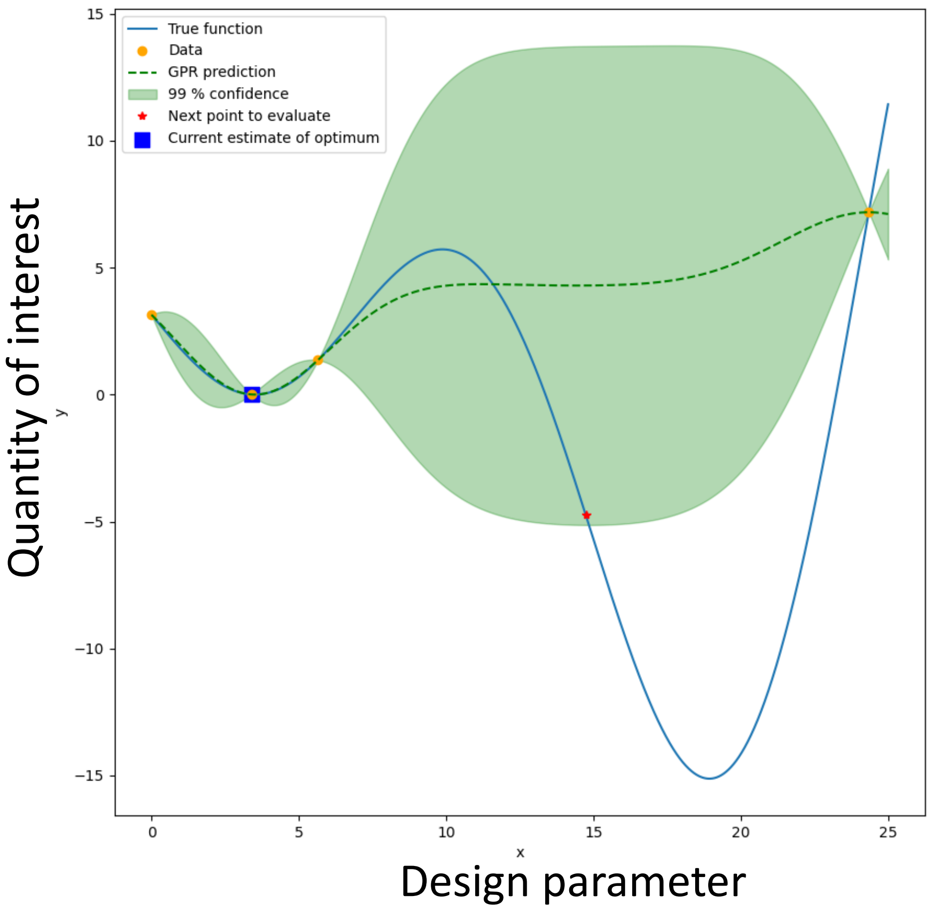
Application: virtual engineering of biofuels

- Process lignocellulose-rich biomass into biofuel
- Inputs: ~10 chemical and processing design parameters
- Large design space to search for optimal chemical process



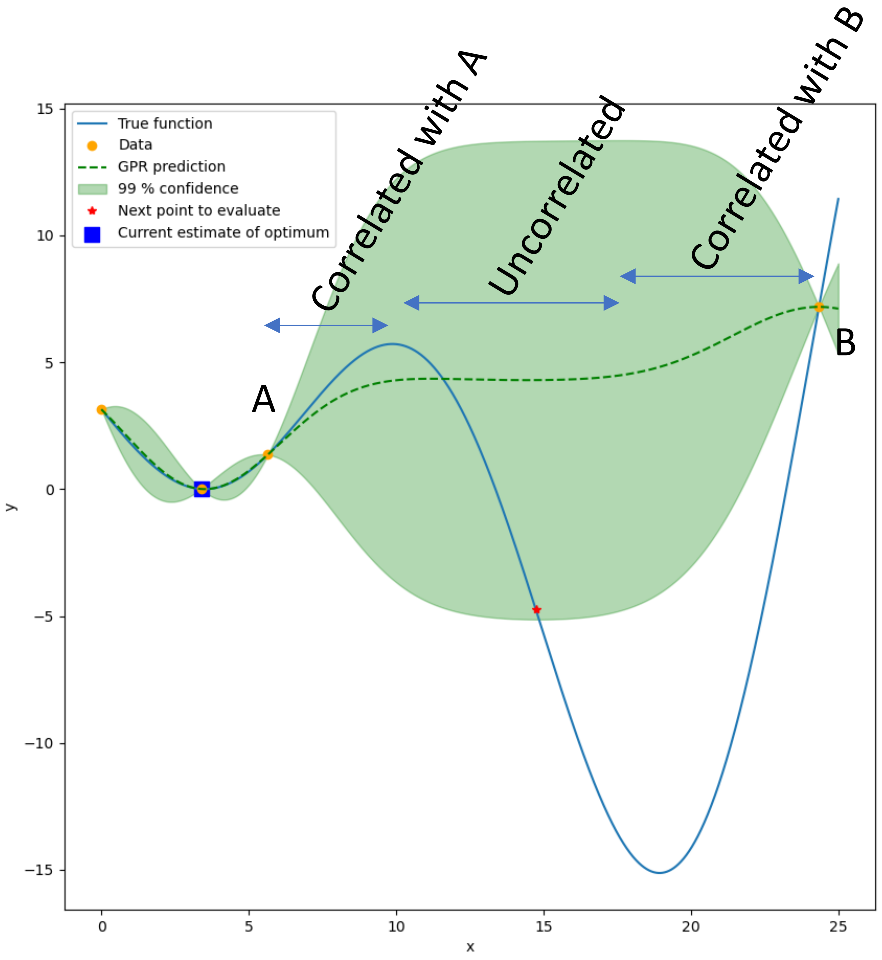
Framing the optimization problem: Single fidelity models

- Quantity of interest
  - E.g., $ cost of generating one unit of biofuel
- Design parameter
  - E.g., temperature of the reactants
- Constraints
  - Fixed computational budget
  - Acceptable ranges for design parameters
- Samples require expensive simulations
- Adaptive computing
  - Given the results of previous simulations, which point in the design space should we simulate next?



Gaussian processes for surrogate modeling

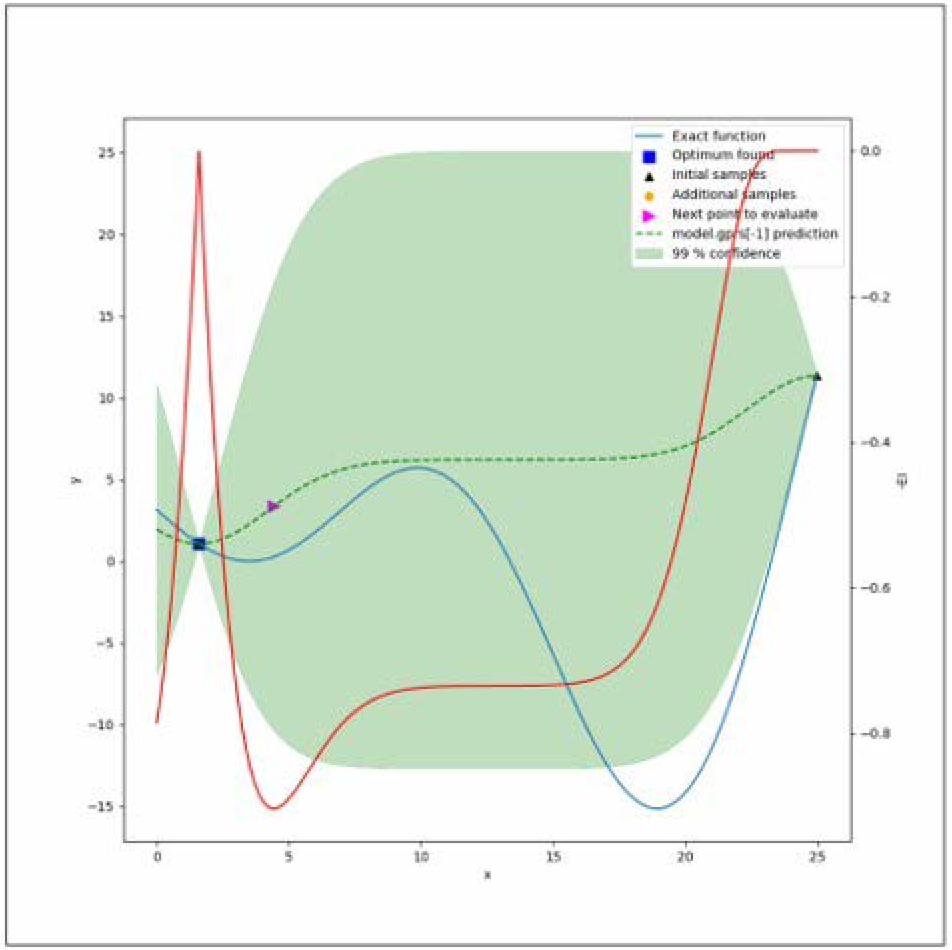
- A way to do interpolation
  - Spline interpolation is a special case of GP models
- Provide an estimate of uncertainty
  - Estimated variance is related to the number of samples and the smoothness of these observations
- Analytical representation can be differentiated for optimization



Bayesian optimization

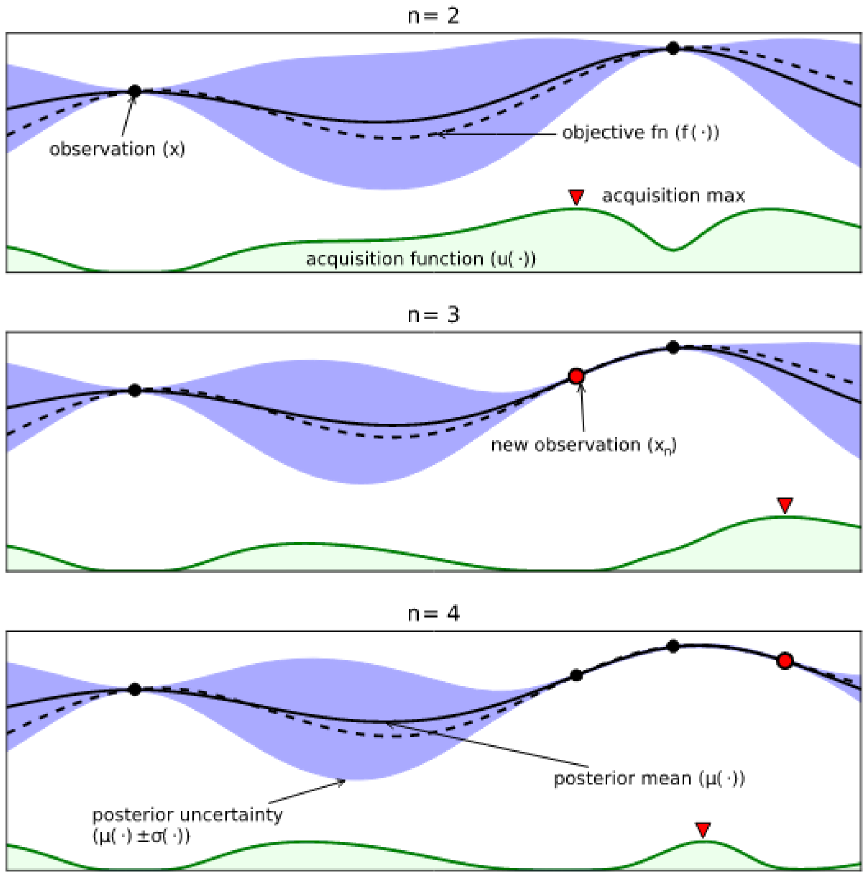
- Avoiding brute force parameter sweep
- Algorithm
    1. Run some trial simulations (initial random samples)
    2. Train a GP model using simulation data
    3. Minimize acquisition function to find next sample point
    4. Run a simulation at this sample point
    5. Repeat steps 2-5

    Animation ->



Acquisition function (AF)

- Key ingredient for automatic model training
  - AF minimum is the optimal place to sample next (not necessarily obj func minimum)
- If goal is to develop a globally accurate surrogate,
  - AF could be the variance
- If the goal is to find the global minimum,
  - AF could be Expected Improvement, an algorithm that balances exploitation and exploration [1]



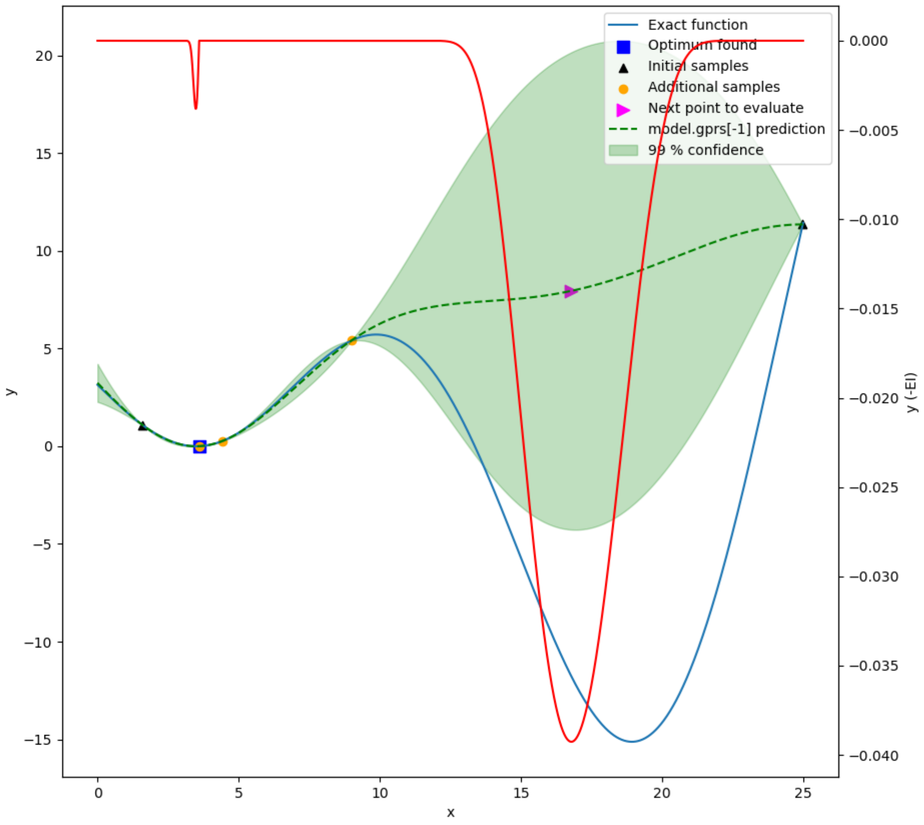
Shahriari et al. *Proc. IEEE* 2016

$$AcqFunc(x) = E[\max(Y(x) - f_{max}, 0)]$$

[1] Jones et al., *Journal of Global Optimization* 1998

Expected improvement and global minimization

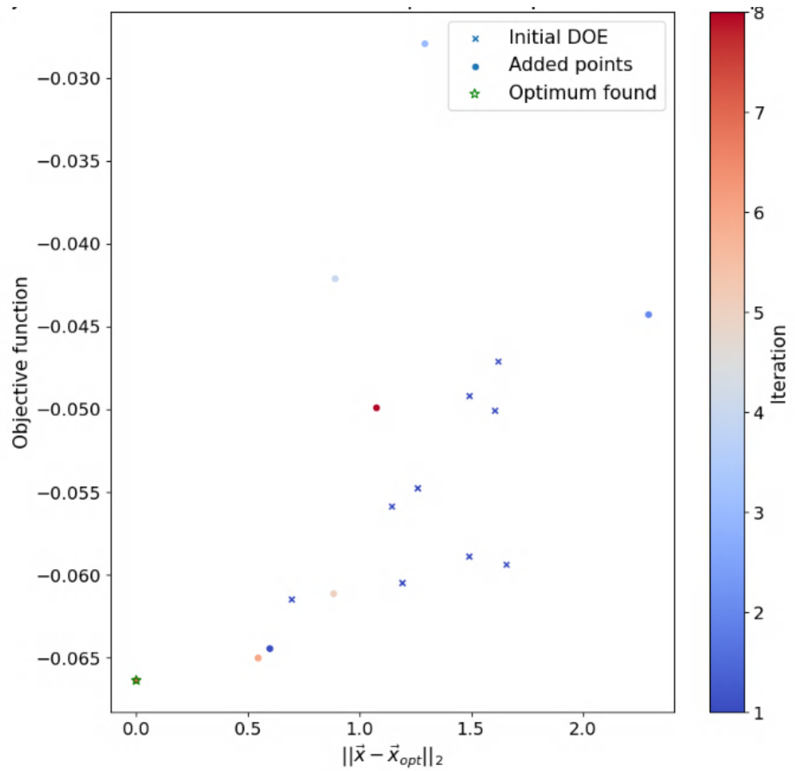
Expected improvement $-EI(x) = E[\max(f_{min} - Y(x), 0)]$



Virtual engineering of biofuels

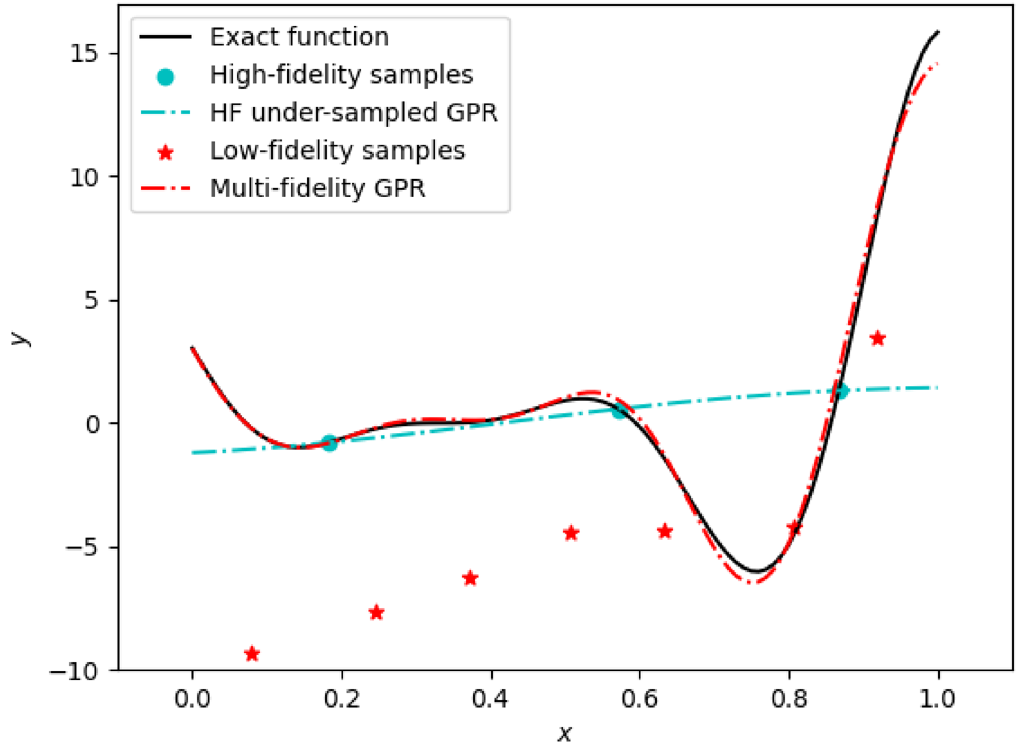
- Maximize O2 uptake rate
- 8 design parameters:

Parameter name	VE default	Min	Max	Final
Fraction of solids that is xylan	0.263	0	1	0.32
Fraction of solids that is glucan	0.4	0	1	0.29
Porous fraction of the biomass particles	0.8	0	1	0.64
Initial concentration of acid	1e-4	0	1e-3 (1)	1e-3
Steam temperature (C)	150	3.8	250.3	170
Fraction of insoluble solids	0.745	0	0.99 (1)	0.99
Enzymatic load	30	0	1000	57
FIS_0 target	0.05	0.005 (0)	1	0.005



Multi-fidelity Gaussian Process

- Multi-fidelity model assumes:
  - Correlation of low- and high-fidelity models
  - $y_{MF} = y_{LF}\rho[x] + \delta[x]$
  - $\rho[x]$ and $\delta[x]$ are low-order polynomials called bridge functions
- Algorithm
  - Given high-fidelity y_{HF} and low-fidelity y_{LF} samples
  - Solve the least squares problem $y_{HF} \approx y_{LF}\rho[x] + \delta[x]$ for the polynomial coefficients

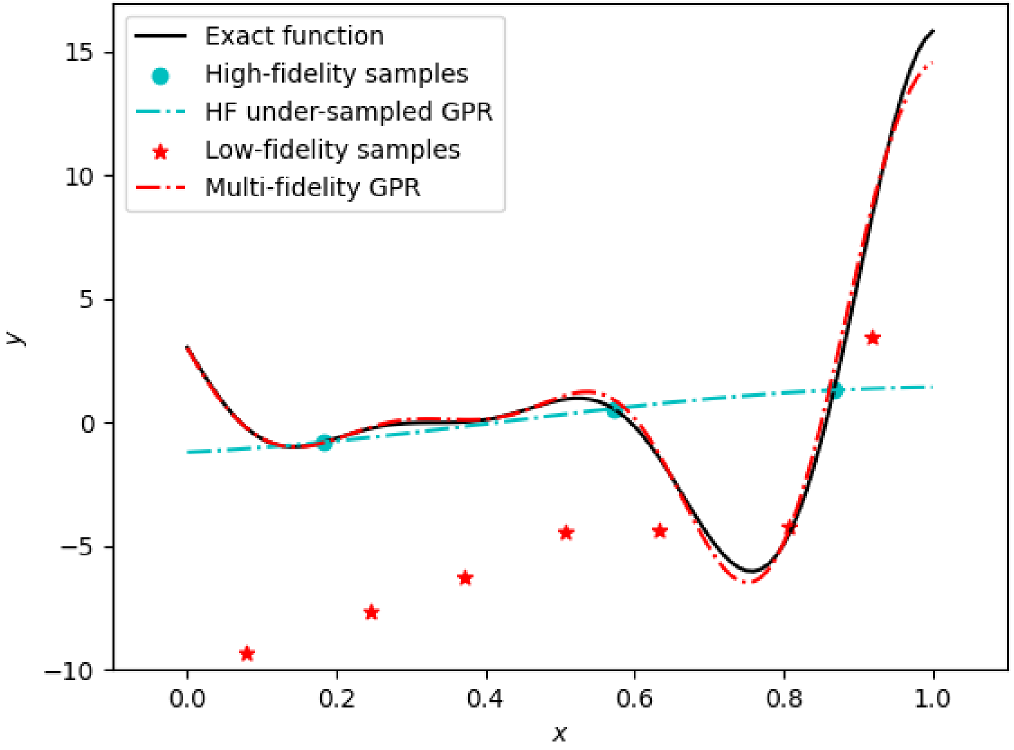


Multi-fidelity optimization

- Acquisition function written for bi-fidelity GP determines which x to sample next
- How to decide from which fidelity level to sample from?
  - Multi-fidelity acquisition function is weighted by computational cost

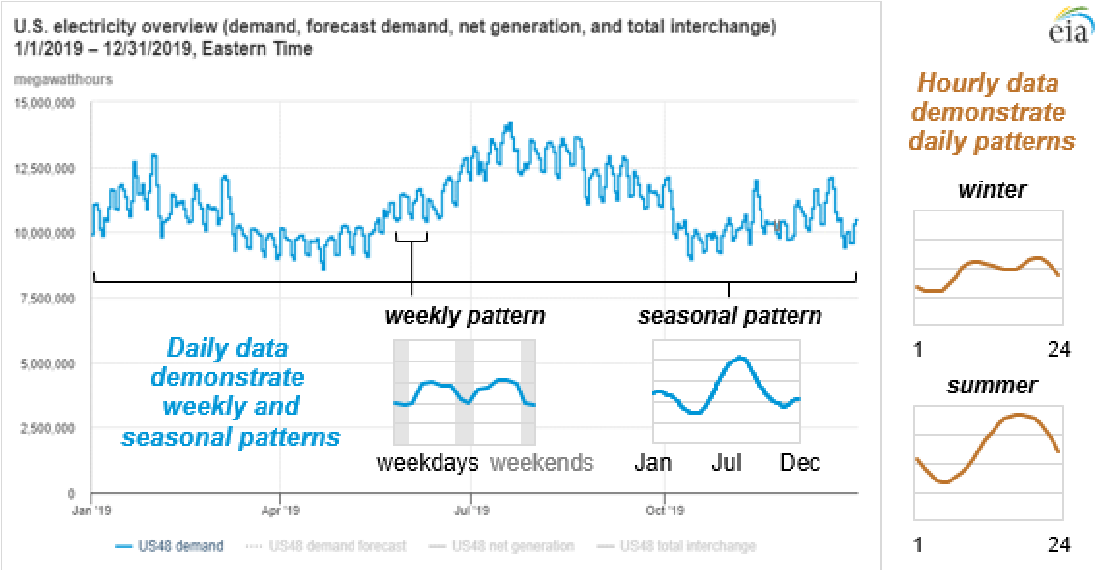
$$\max\left(\frac{EI_L}{C_L}, \frac{EI_H}{C_H}, \ldots\right)$$

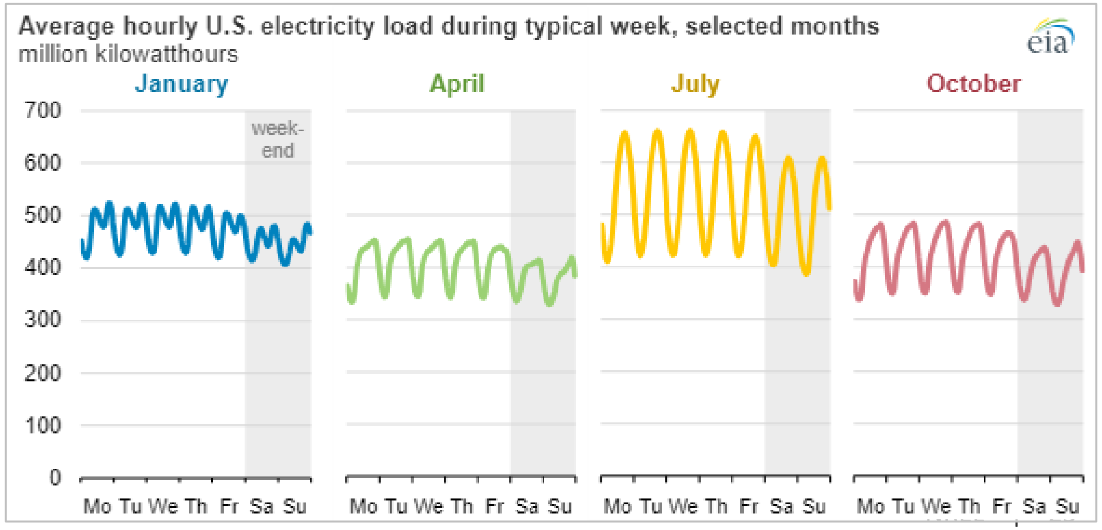
- Janelle Domantay is working on more sophisticated criteria



Why adaptive computing for buildings?

- The total load seen by utilities is the aggregate of individual loads
  - Buildings: residential, commercial, and industrial
- To decarbonize, we
  - Electrify energy consumption: demand goes up
  - Replace generation with clean sources: fluctuating generation
- The buildout of a 100% decarbonized grid is prohibitively expensive
  - NREL research shows that the last ~10% is economically infeasible
  - Alternate solutions, particularly controlling loads will play a major role

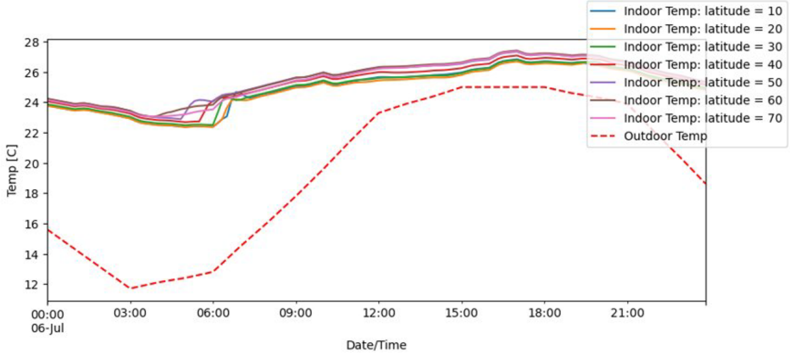




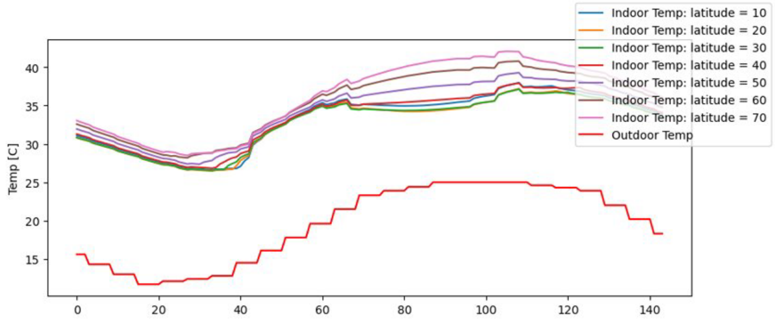
Application to energy use of buildings

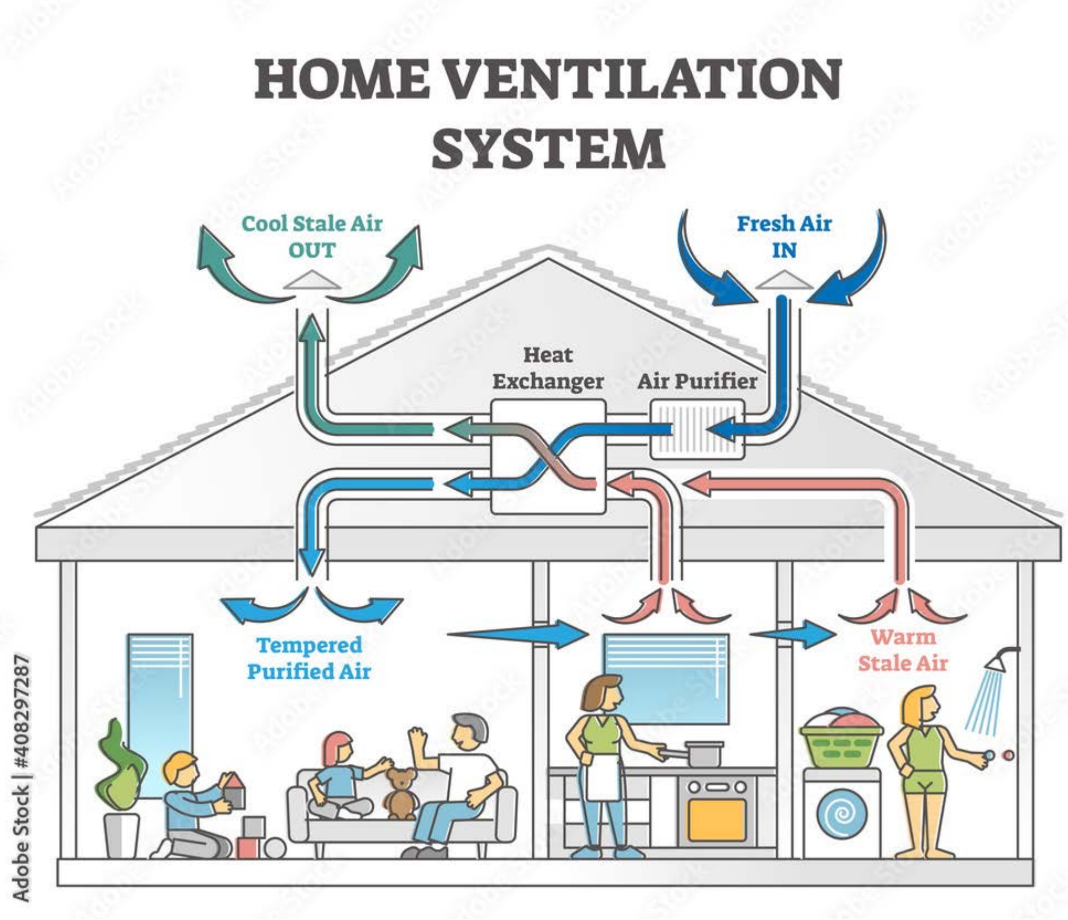
- Two fidelity models for building HVAC
- Low fidelity model inaccurate at high latitudes, but captures the trends

High fidelity



Low fidelity

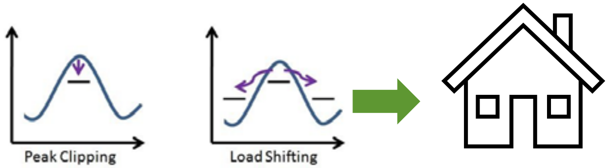




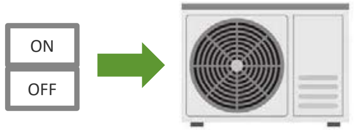
Adaptive computing serves the various scales needed

- For monthly, weekly, and day-ahead projections
  - Tools like ***EnergyPlus*** help evaluate large scale scenarios factoring in large **uncertainties of weather, social events, and DER behavior**
  - Explore/optimize pathways of orchestrating **control and coordination**
  - Provide **highest fidelity** in modeling building physics
  - Optimization can become **computationally expensive**

- For hourly and time of day projections
  - Tools like ***OCHRE*** can 'simplify' assumptions while still within bounds defined by EnergyPlus
  - Be more responsive to grid conditions



- For immediate, local control of equipment
  - **Reduced order surrogates** can be very responsive, use little compute while living at the edge, and deliver targeted control within buildings
  - Can be responsive to larger 'supervisory' guidance/commands from OCHRE or EnergyPlus to relax or tighten control execution to meet the larger community scale energy goals

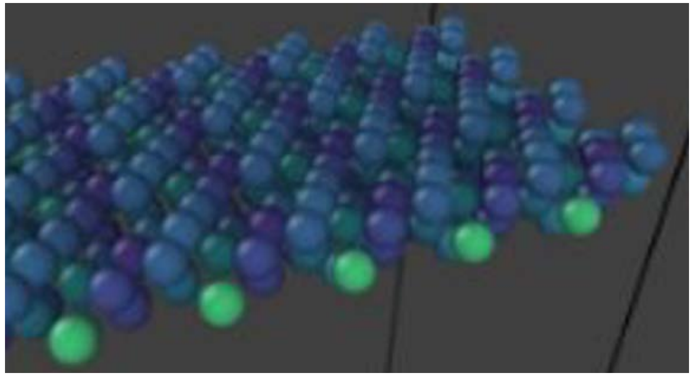


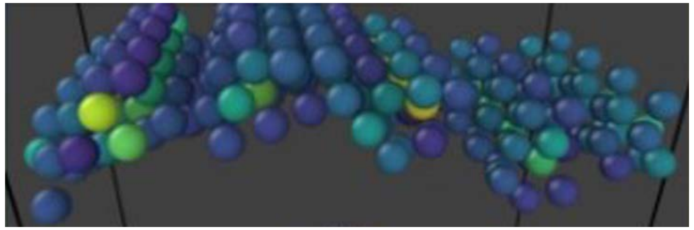
Connection to ARIES

- The adaptive computing framework supports the buildout of the ARIES Virtual Emulation Environment
  - There is a need to represent local to regional scale energy systems and buildings contribute an oversized share to the total picture
  - Communities that use ARIES have questions that require an adequate representation of their buildings stock
  - Many of these questions require the evaluation of technology choices involving DERs and controllability of these devices, individually and at scale
  - Multi-fidelity capabilities are needed to emulate and validate the solutions

Application to material synthesis for solar PV

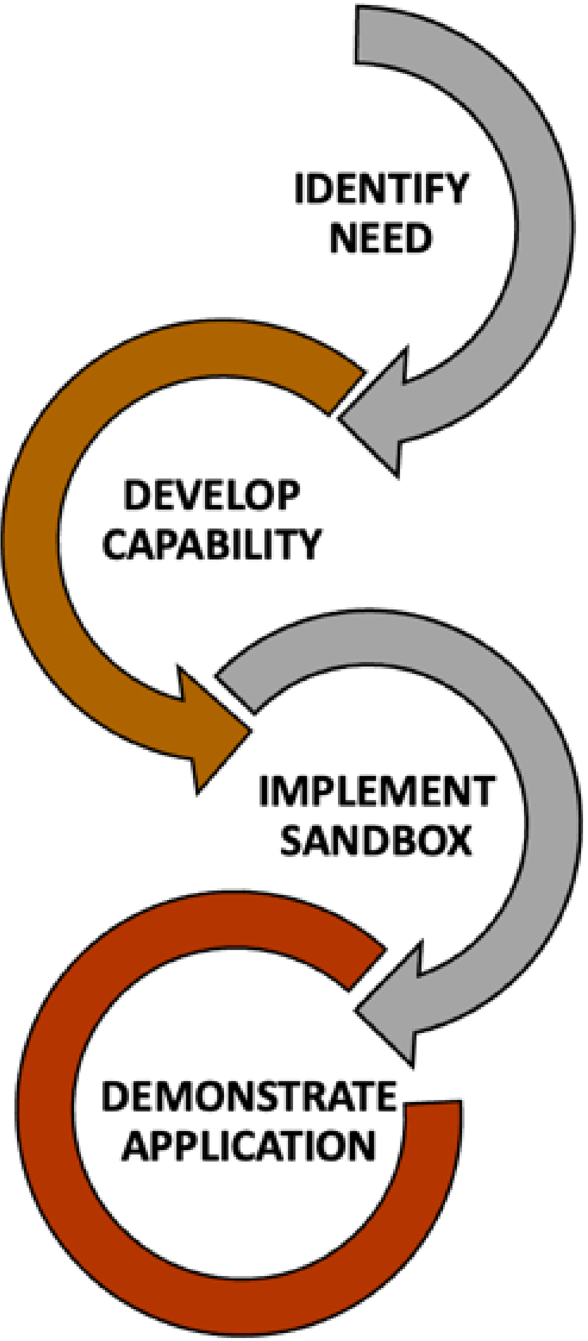
- Objective- optimal atomic structure to maximize PV performance
- Design parameters- synthesis gas composition, temperature, and pressure
- Fidelity levels
  - LF- Molecular dynamics (interatomic potentials are uncertain)
  - HF- **Automated experiment**





Adaptive Computing Project Staffing

1. AC infrastructure, surrogate model management (AC Leads: Kevin Griffin, Ryan King)
2. Optimization, active learning, UQ (AC Lead: Juli Mueller)
3. Engagement of applications that guide development of AC infrastructure
   - Power grid stability with renewable energy sources (AC Leads: Jibo Sanyal, Deepthi Vaidhynathan)
     SME: Jen King, Rob Hovsapian
   - Biofuels Virtual Engineering (AC Leads: Marc Day, Kevin Griffin)
     SME: Nicholas Carlson (TEA), Andrew Glaws (Surrogate models), Hari Sitaraman (High-fidelity simulation), Ethan Young + Olga Doronina (Optimization + Workflow)
   - Multiscale Biomass Modeling (AC Lead: Hilary Egan)
     SME: Peter Ciecielski
   - Virtual Material Synthesis (AC Lead: Hilary Egan)
     SME: Garritt Tucker (Mines), Andriy Zakutayev
   - Vapor Deposition for Halide Perovskites (AC Lead: Marc Henry de Frahan)
     SME: Dave Moore
   - Catalyzed polymer upcycling (AC Lead: Bruce Perry)
     SME: Matt Carbone (BNL), Mike Crowley



Full project scope (FY23-FY25)

- Understand the needs and challenges of scale up across applications
  - Data and model inventory
- Develop resource management tools and base capability
  - Model (re-)design, cross-fidelity model management, feedback loop
- Application integration and testing on readily available models and data
- Demonstration across diverse applications, results publication
  - Further tuning and feature expansion

Thank You

www.nrel.gov

NREL/PR-2C00-88097

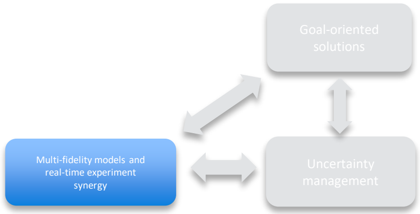
This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Photo from iStock-627281636

NREL

Transforming ENERGY

Multi-fidelity Modeling

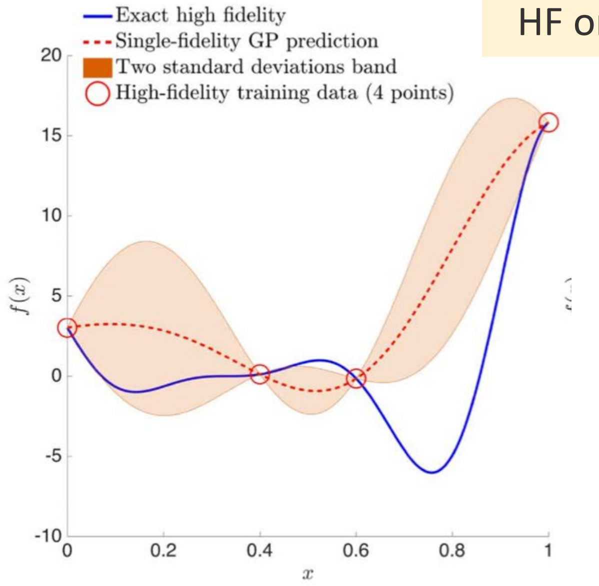


Most applications feature an assortment of models of widely varying fidelities, developed for different purposes:

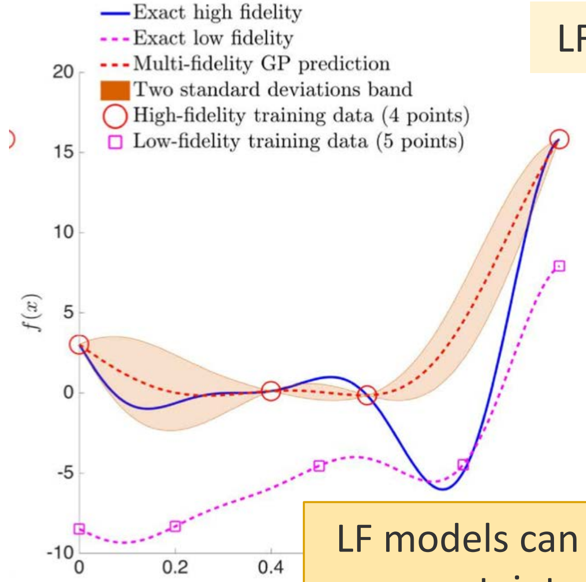
- Experiment: “Truth”, but limited operational regime
- High-fidelity: Physics-based (PDE/ODE), costly
- …
- Low fidelity surrogates: data-driven AI/ML, PINNs, …

1. Can we orchestrate levels, leveraging the appropriate hardware, to maximize bang for our computational buck?
2. How do domain-specific models and goals impact this?

HF only



LF+HF



LF models can reduce uncertainty with minimal extra cost

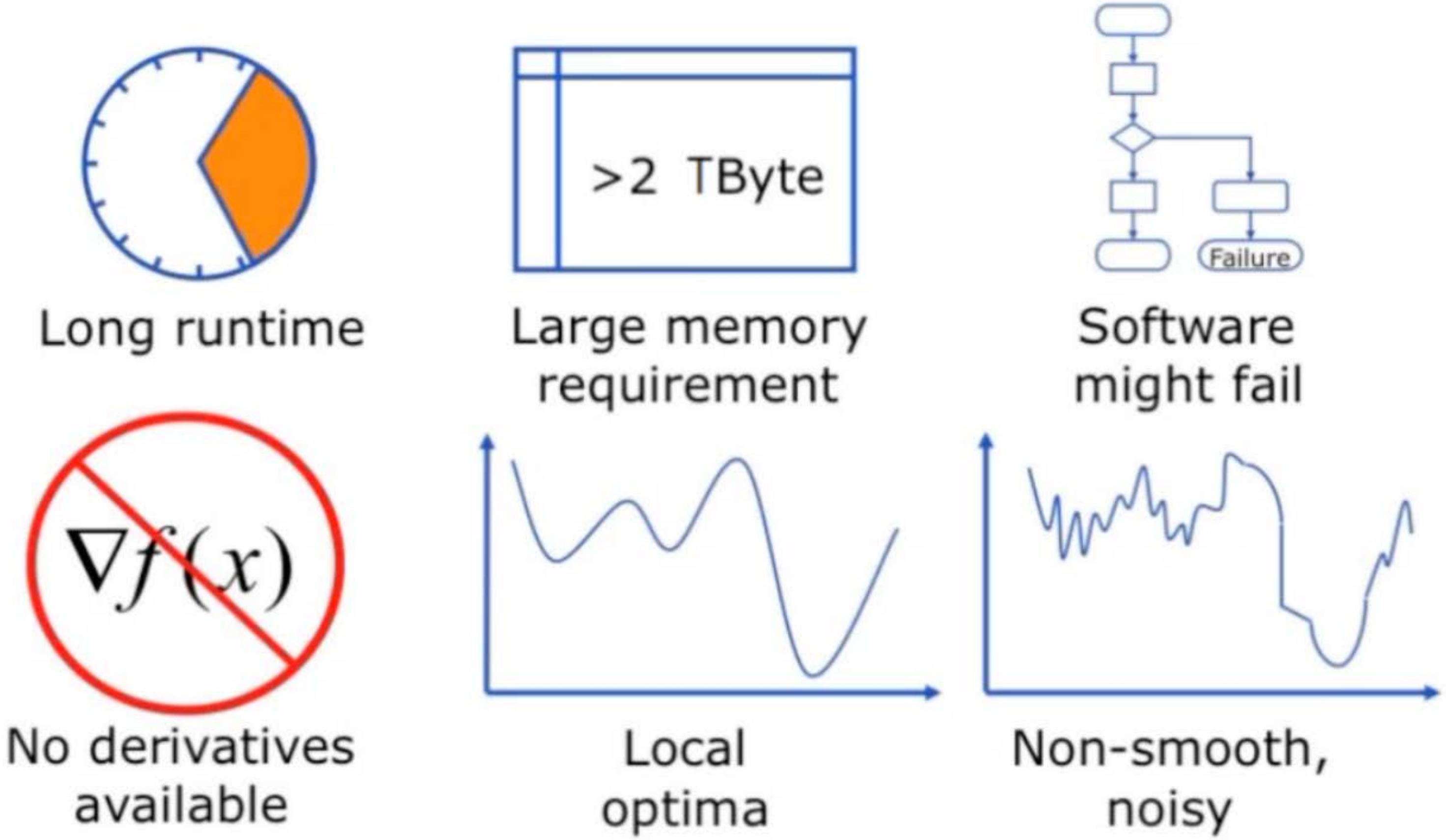
Typical MF control variate estimator:

$$\mathrm{Var}\left(\widehat{Q}^{HF,CV}\right) = \mathrm{Var}\left(\widehat{Q}^{HF}\right)\left(1 - \frac{r}{1+r}\rho^2\right)$$

Idea: introduce and train new ML model such that the variance of the new MF model is minimized

$$\min_{\boldsymbol{\gamma},\boldsymbol{\theta}} \quad \mathrm{Var}\left(\widehat{Q}^{HF}\right)\left[1 - \frac{r(\boldsymbol{\gamma},\boldsymbol{\theta})}{1+r(\boldsymbol{\gamma},\boldsymbol{\theta})}\rho(\boldsymbol{\gamma},\boldsymbol{\theta})^2\right]$$
$$\text{s.t.} \quad C^{tot} = C^{HF}N^{HF} + C^{SciML}(\boldsymbol{\gamma})\left[1 + r(\boldsymbol{\gamma},\boldsymbol{\theta})\right]N^{HF} + C^{train,SciML}(\boldsymbol{\gamma})$$

Black-box optimization



Gaussian process model

- Wang (2020) "An Intuitive Tutorial to Gaussian Processes Regression."
- Two random (normally distributed) uncorrelated processes:

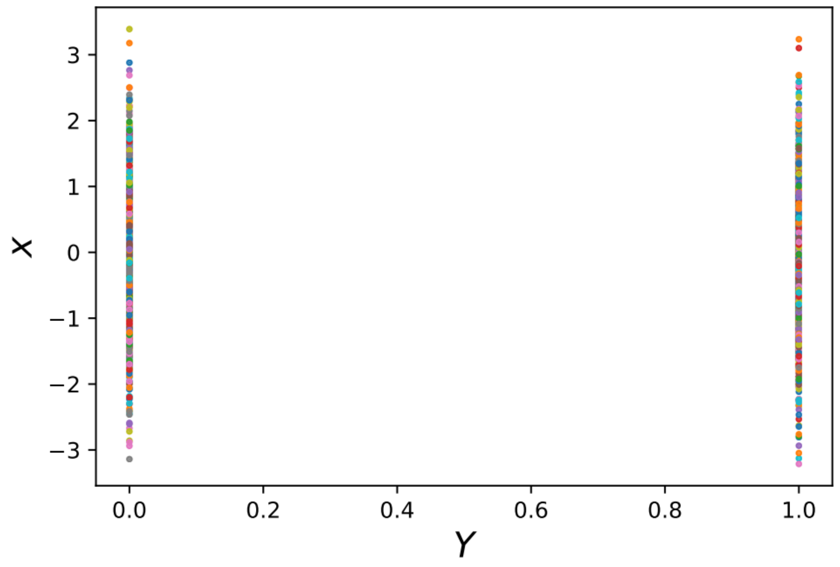
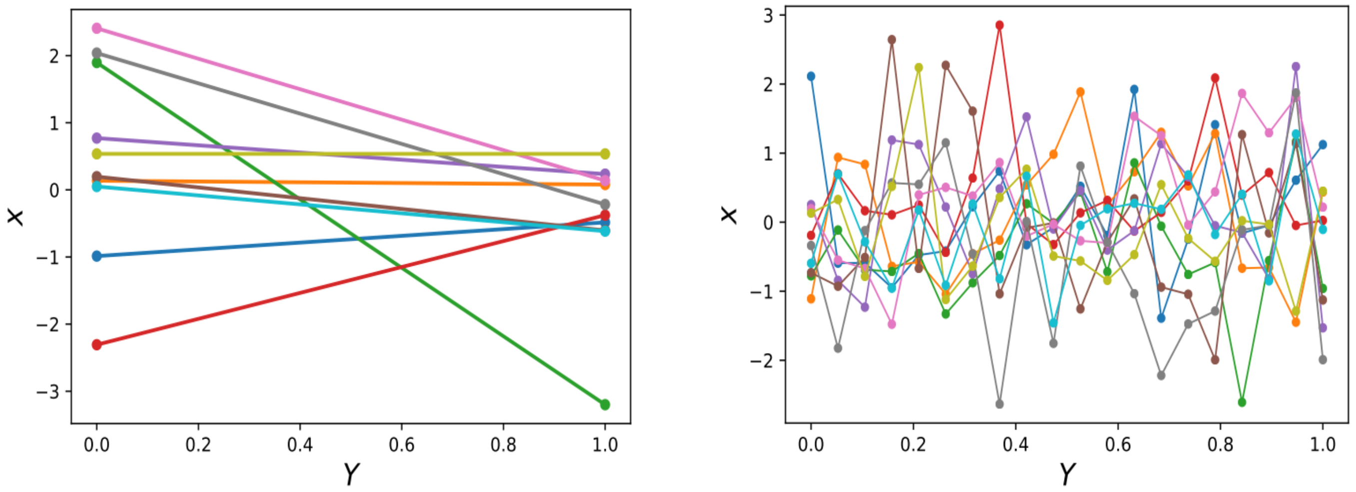


Figure 3: Two independent uni-variate Gaussian vector points were plotted vertically in the Y, x coordinates space.

Adding more uncorrelated processes

- E.g., cost of a unit of biofuel versus a butterfly’s wing position



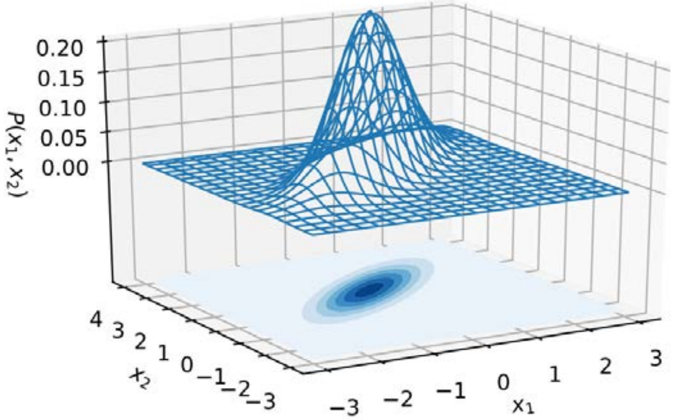
(a) Two Gaussian vectors

(b) Twenty Gaussian vectors

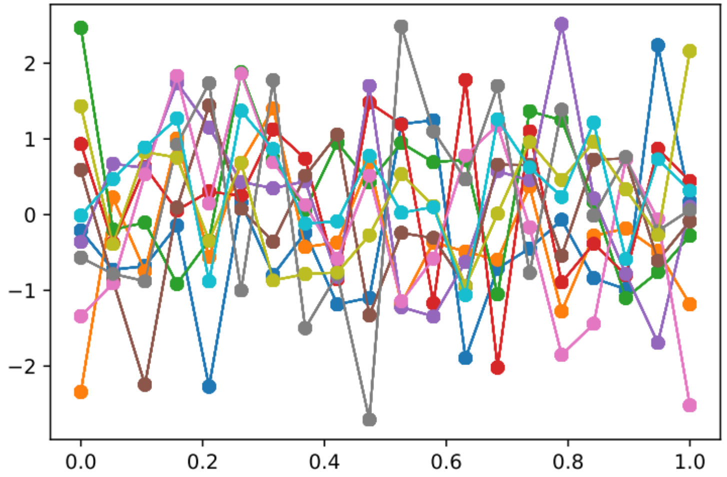
Figure 4: Connecting points of independent Gaussian vectors by lines: (a) Ten random selected points in two vector x_1 and x_2 , (b) Ten random selected points in twenty vectors $x_1, x_2, \ldots, x_{20}$.

Assume the samples are correlated

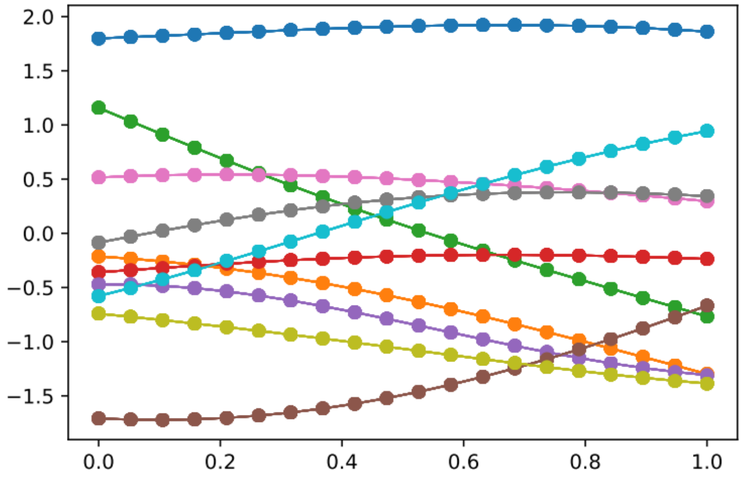
- E.g., cost of a unit of biofuel versus reactant temperature



(a) 3-d bell curve

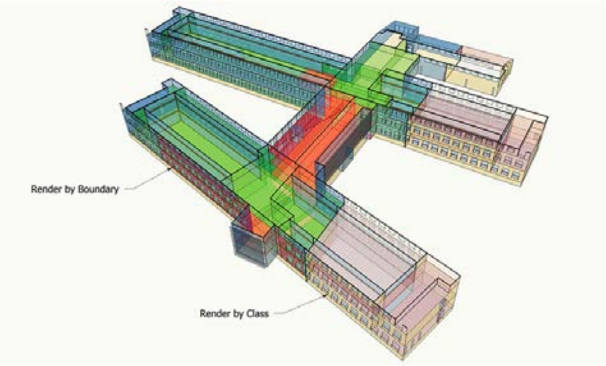


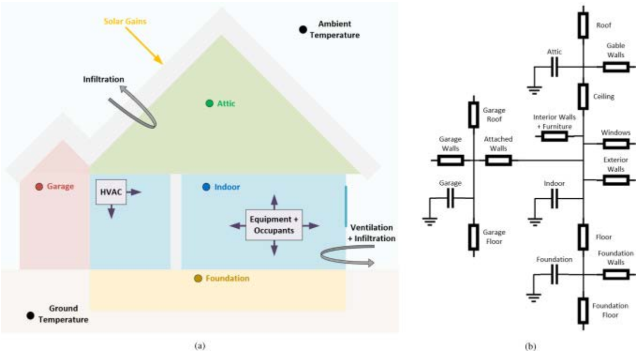
(a) Ten samples of the 20-VN prior with an identity kernel



(b) Ten samples of the 20-VN prior with a RBF kernel

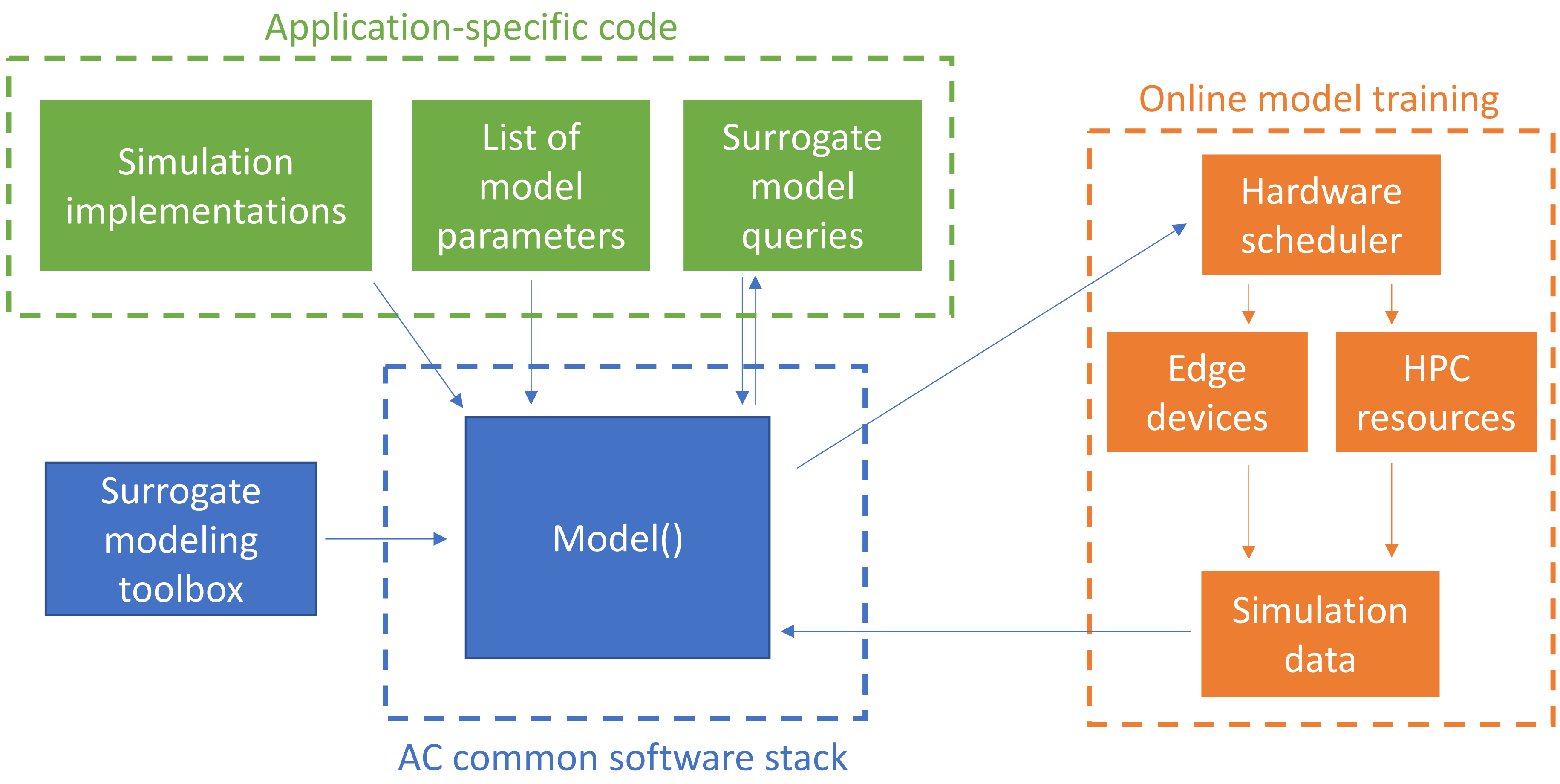
Different fidelity regimes for building mod, sim, & control



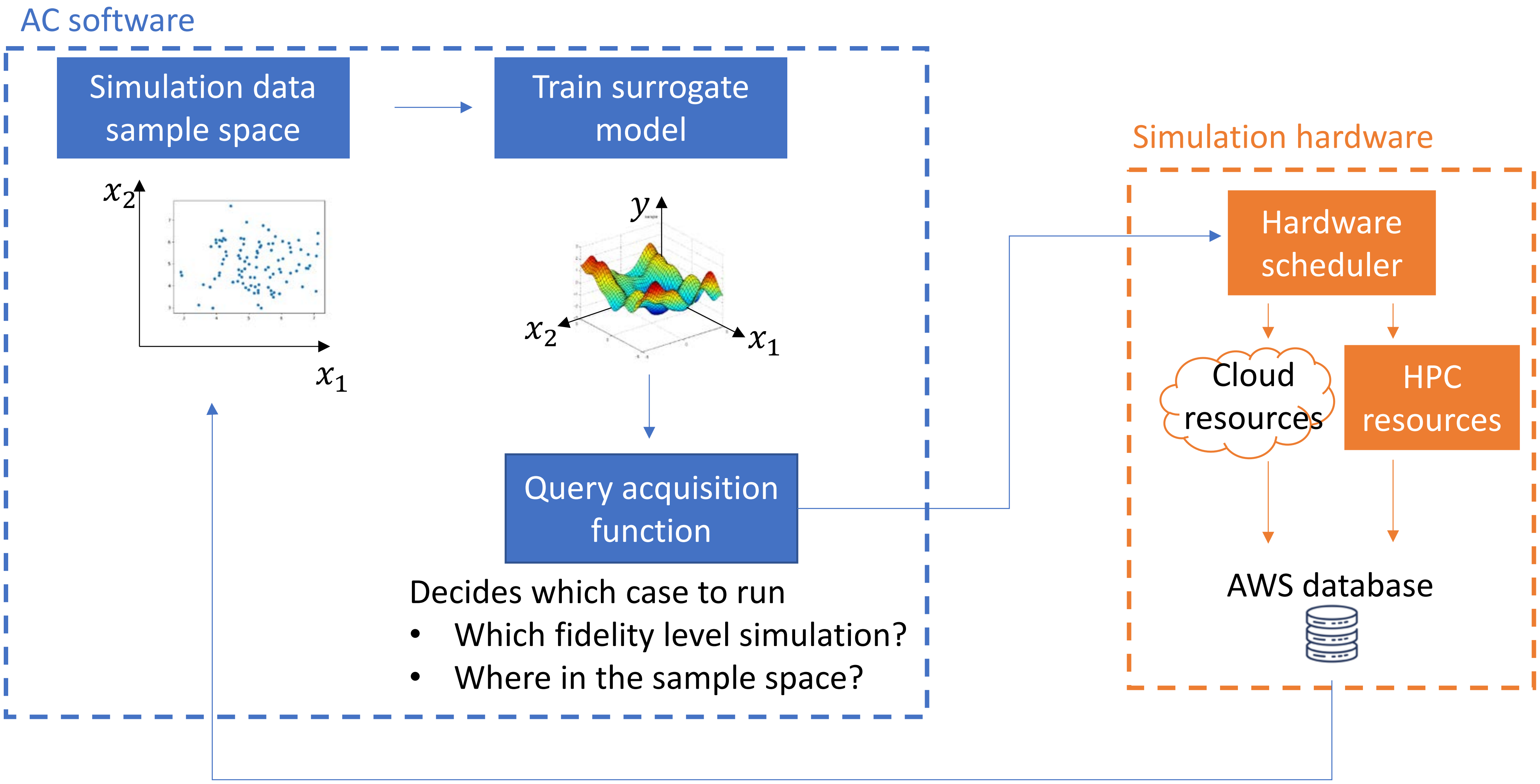


	EnergyPlus	OCHRE	Reduced order model
Applications	**Modeling building** for residential and commercial buildings	**Building to grid integration** for residential single zone buildings	**Building to grid integration** for commercial and residential single zone buildings **control applications**
Fidelity	High/Ground truth	Medium	Low / grey box model
Computational Resources-need	High	Medium	Low / dependent on data used to train the model.
Inputs	Detailed building geometry, weather files, detailed equipment information, schedules, occupancy information and models.	Building geometry, weather files, equipment information, occupant profile and schedules.	Time series data of indoor air temperature, cooling/heating effort, loads, outdoor air weather conditions.
Outputs	Building energy consumption, sensors and building physics data.	Building energy consumption, sensors and building physics data	Room air temperature, Cooling/Heating Power
Adaptive computing	Develop multi-fidelity and goal-oriented building energy models that spans across spatial and temporal scales to support clean energy transition of communities at scale.		

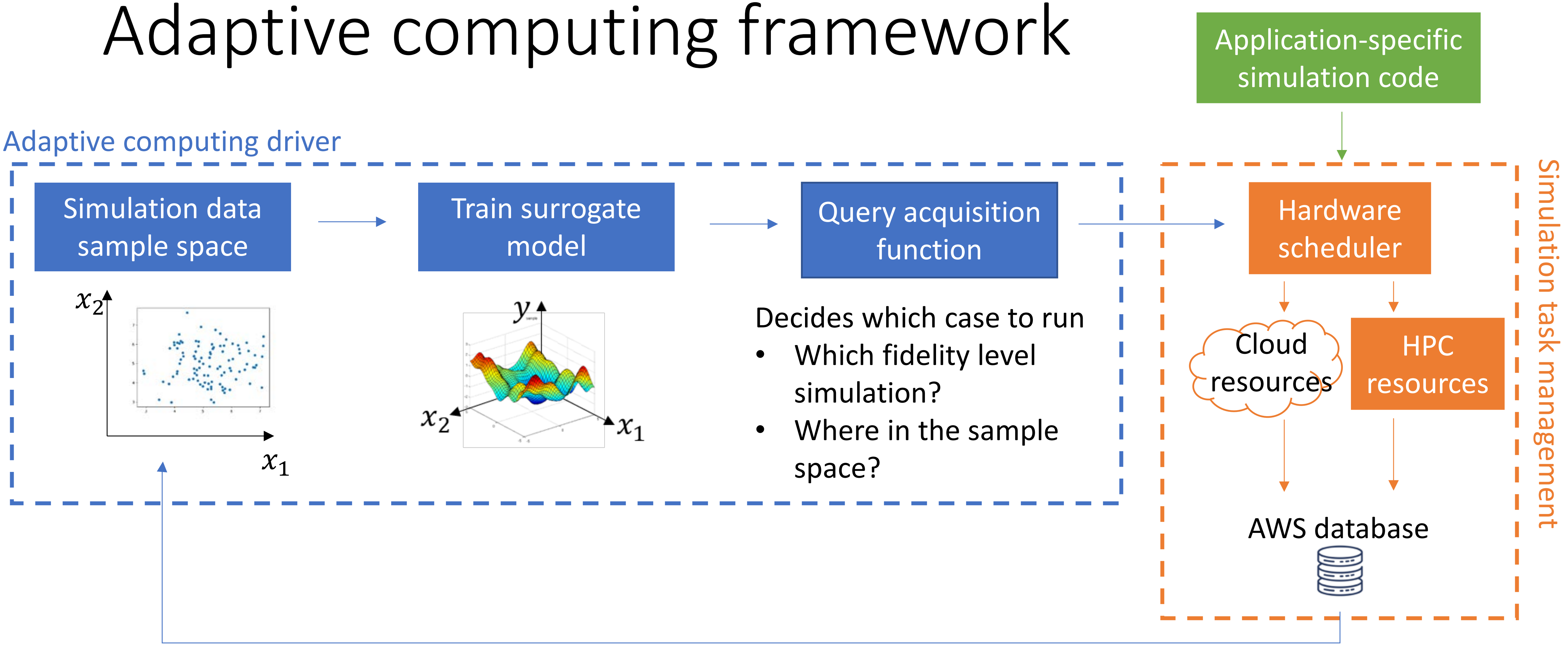
Application-flexible interface

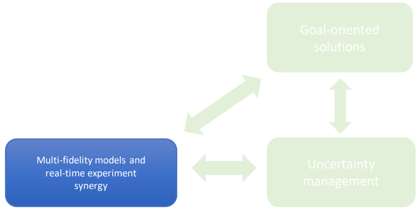


Adaptive computing framework



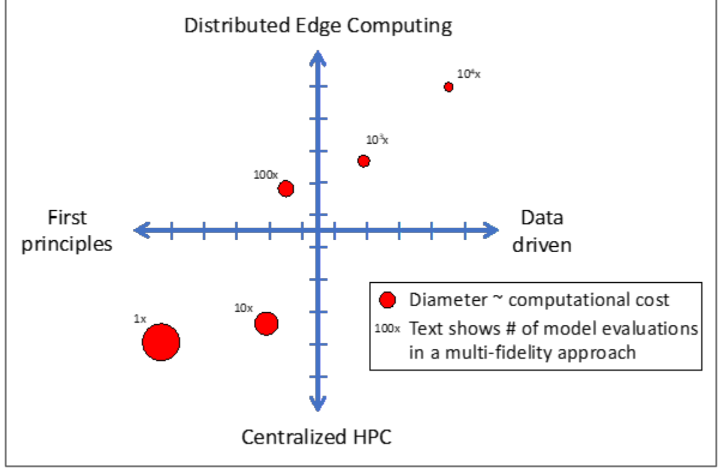
Adaptive computing framework

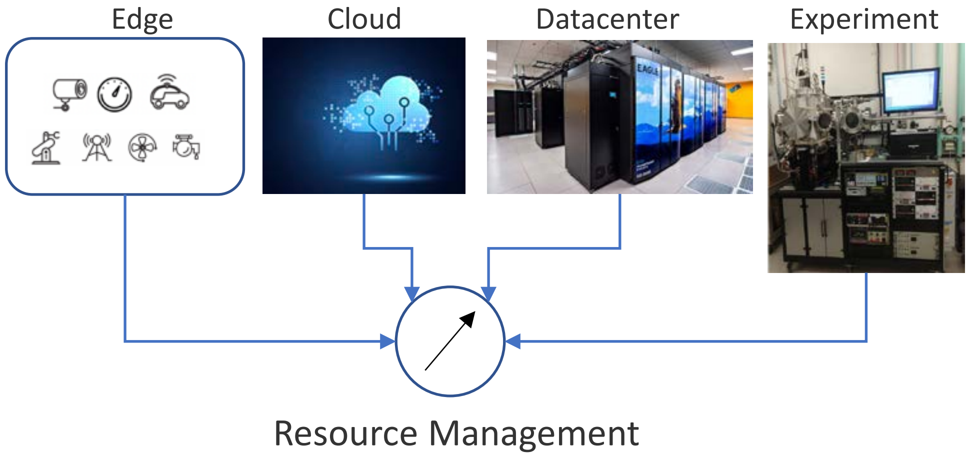




Model and “experiment” (physical or hi-fidelity) synergy

Model/resource matching requirements





Sample point requester:
- How many new samples?
- What fidelity level of new samples?
- Experiment/high-fidelity or surrogate?

Constrained muti-fidelity optimization: Maximize a utility function such that compute constraints satisfied

- Return list of experiments / simulations to be carried out
- Higher fidelity results can also to inform / improve lower fidelity surrogates

Requires optimization modeling, sampling strategy development & UQ developments