Hierarchical Models for Insightful Machine Learning

Science Accelerator

Markus Kaiser
www.mrksr.de

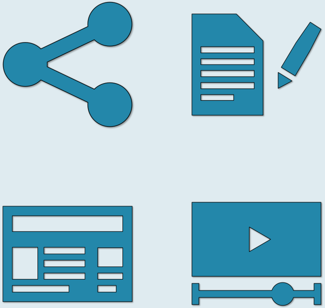
3 February 2021

University of Cambridge, Siemens AG

Real-world machine learning

Human-centered ML

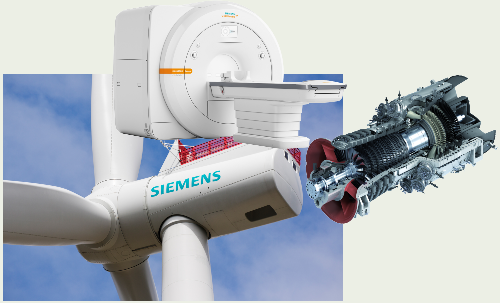
- Confined systems
- (Seemingly) mild consequences



Examples: Games, Search, NLP

Scientific and Industrial ML

- Real-world interaction
- Safety and semantics are critical

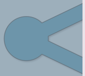


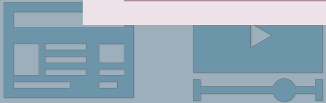
Examples: Machine control, commissioning

Real-world machine learning

Human-centered ML

- Confined systems
- (Seemingly) mild consequences

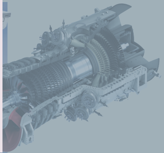




Examples: Games, Search, NLP

Scientific and Industrial ML

- Real-world interaction
- Safety and semantics are critical



Examples: Machine control, commissioning

Properties of real-world ML

- Uncertain models
- Domain knowledge
- Need for trust

The basic ML algorithm

Empirical risk minimization

- Approximate the **global true risk** wrt. loss ℓ

$$\mathrm{R}(f) := \int \ell(f(\mathbf{x}), \mathbf{y})\, \mathrm{p}(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}$$

  with the **local empirical risk** in the available data

$$\mathrm{R}_{\mathrm{emp}}(f) := \frac{1}{N} \sum_{i=1}^{N} \ell(f(\mathbf{x}_i), \mathbf{y}_i)$$

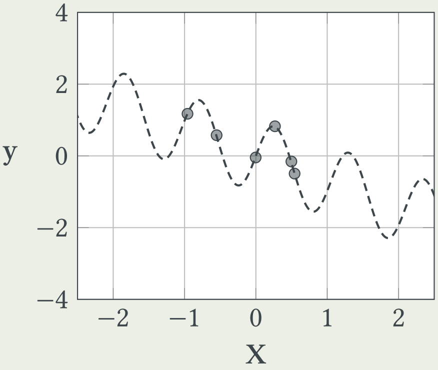
- **Learning algorithm:** Choose a hypothesis space $\mathscr{H} \subseteq \mathscr{F}$ and use

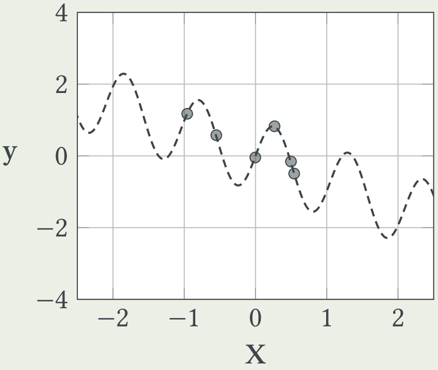
$$\hat{f} \in \operatorname*{argmin}_{f \in \mathscr{H}} \mathrm{R}_{\mathrm{emp}}(f)$$

Generalization is necessary

Scientific and Industrial ML

As data is scarce, **experts** need to tell us how to **generalize aggressively**.

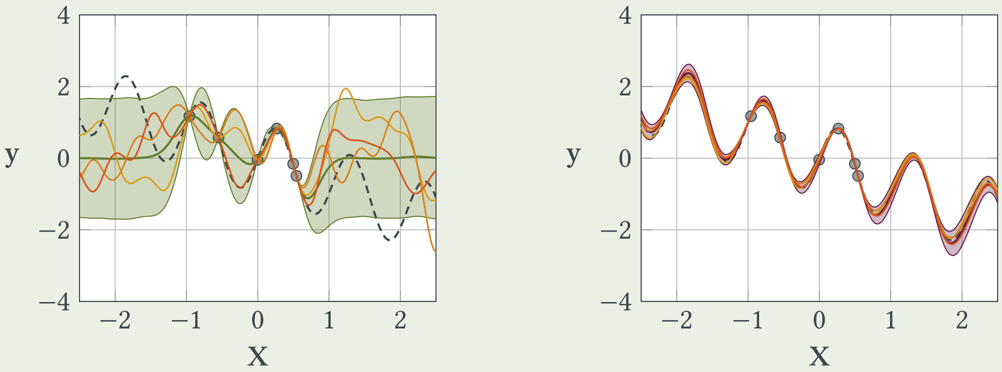




Generalization is necessary

Scientific and Industrial ML

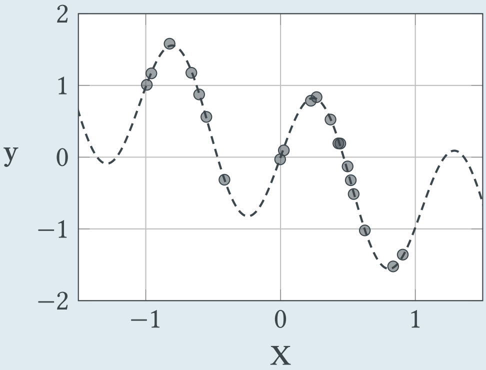
As data is scarce, **experts** need to tell us how to **generalize aggressively**.



Models are known to be imperfect

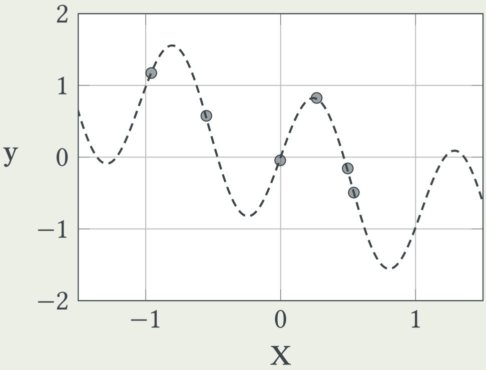
Human-centered ML

- When in doubt, collect more data
- Uncertainties are not so important



Scientific and Industrial ML

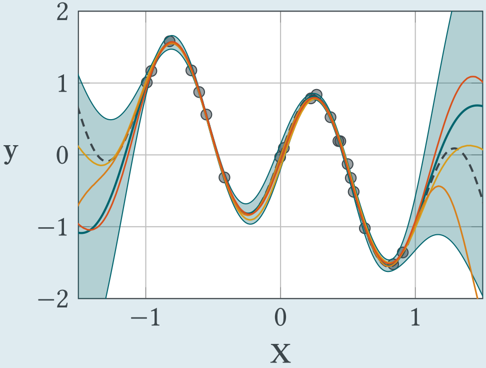
- Need to make do with available data
- Uncertainties are critical



Models are known to be imperfect

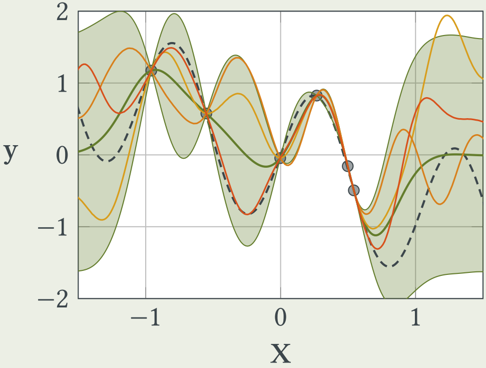
Human-centered ML

- When in doubt, collect more data
- Uncertainties are not so important



Scientific and Industrial ML

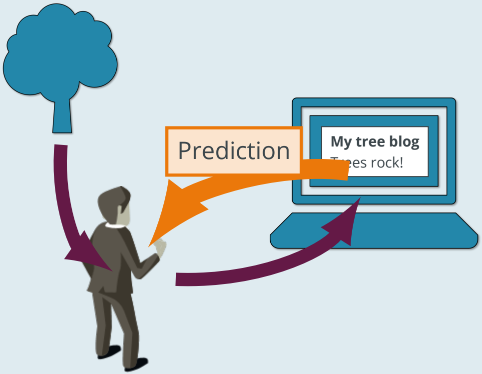
- Need to make do with available data
- Uncertainties are critical



Industry needs interpretability

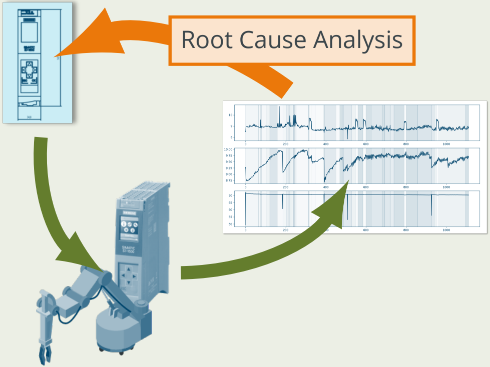
Human-centered ML

- Inform or influence a person
- Understanding is secondary



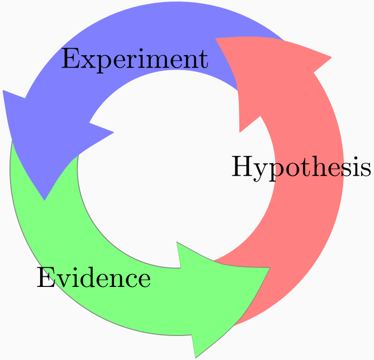
Scientific and Industrial ML

- Create better or safer machines
- Understanding is key



Carl's slide

The Scientific Principle



$$\text{Data} + \text{Model} \overbrace{\rightarrow}^{Compute} \text{Prediction}$$

Knowledge in ML models

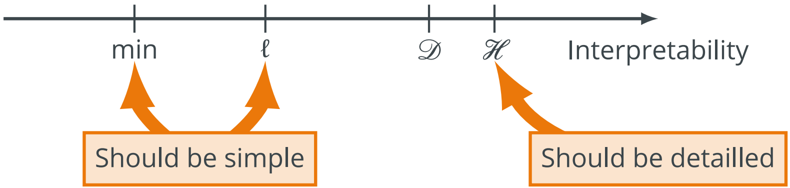
Places to encode knowledge

Observations $\mathscr{D}$	**Data** selection, feature engineering, data augmentation
Hypothesis space $\mathscr{H}$	Choice of **model**, architecture design
Loss function ℓ	Choice of norm, **regularization**
Optimization min	Choice of **optimizer**, initialization, parameter tuning

Knowledge in ML models

Places to encode knowledge

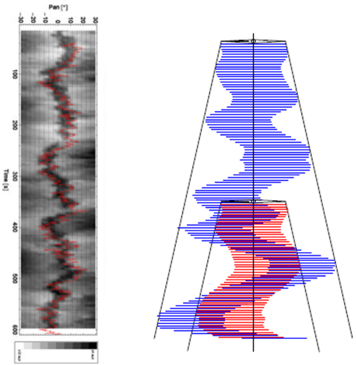
Observations $\mathscr{D}$	**Data** selection, feature engineering, data augmentation
Hypothesis space $\mathscr{H}$	Choice of **model**, architecture design
Loss function ℓ	Choice of norm, **regularization**
Optimization min	Choice of **optimizer**, initialization, parameter tuning



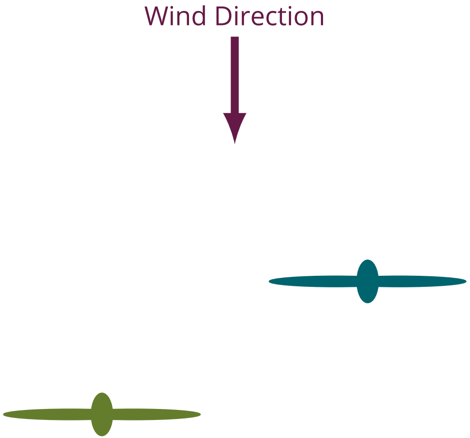
Lillgrund wind farm

Wind and wake propagation

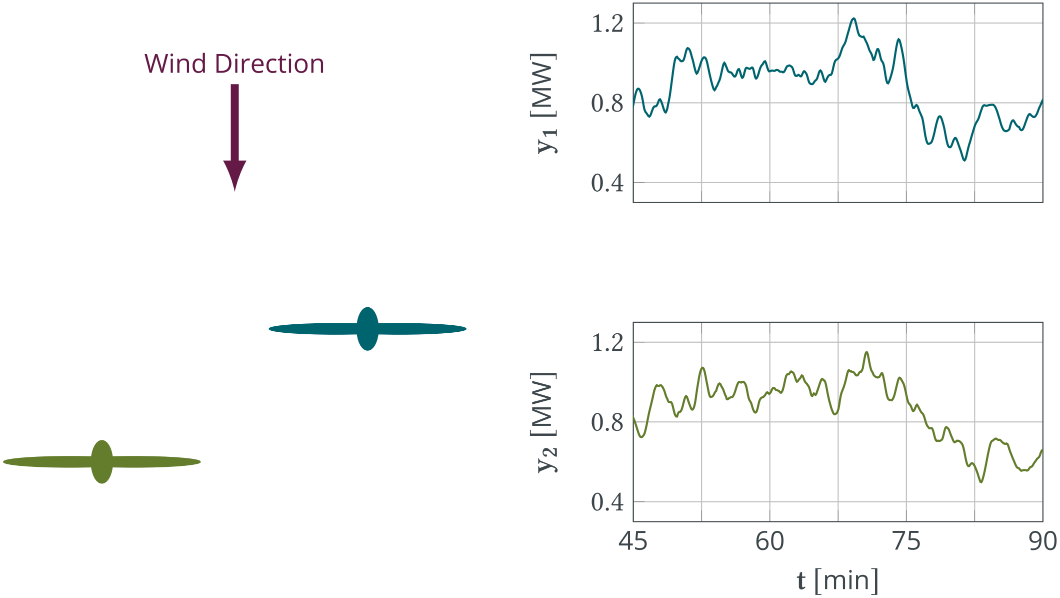
T.J. Larsen et al.: Dynamic Wake Meander Model, Wind Energy (2012)



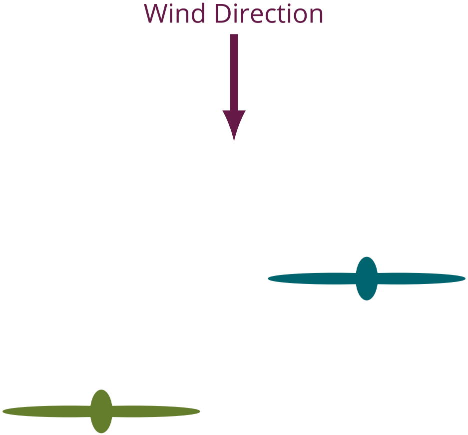
Real-world data

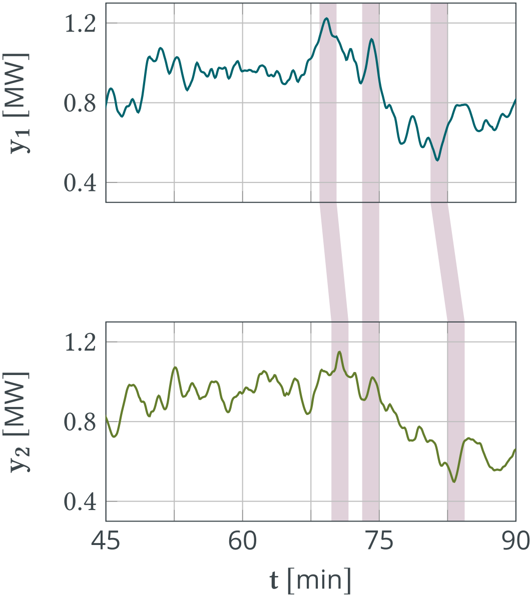


Real-world data

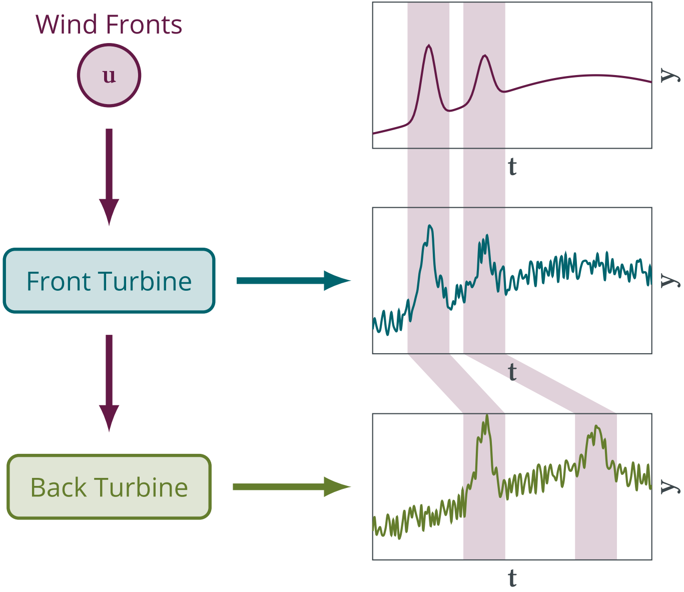


Real-world data

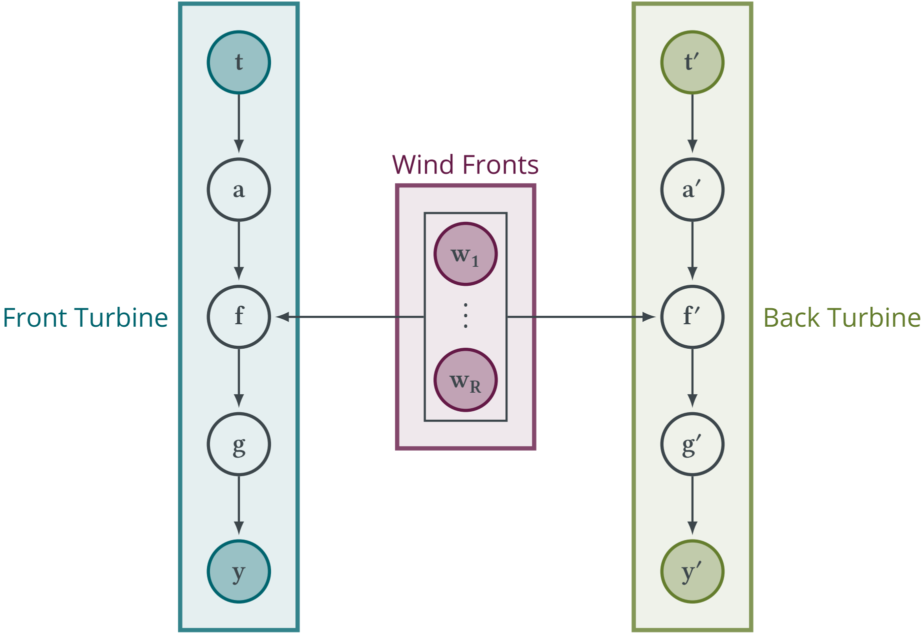




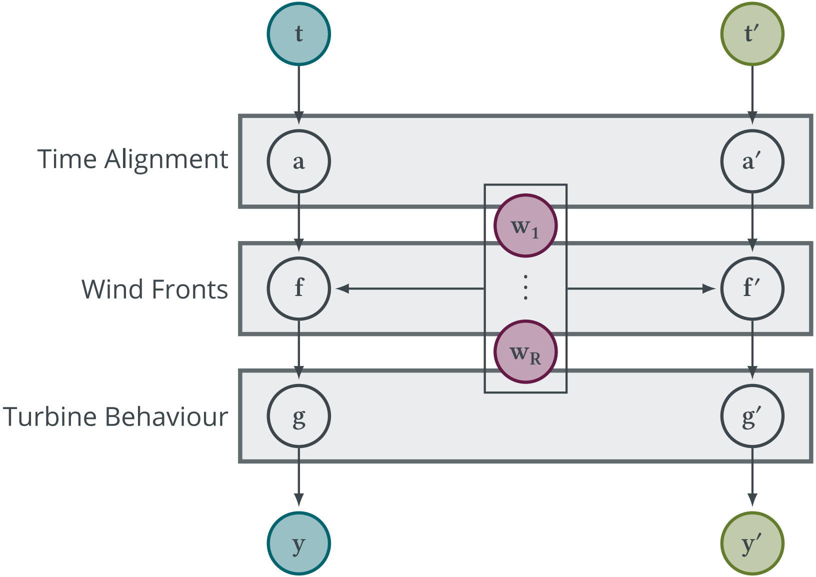
Modelling wind propagation



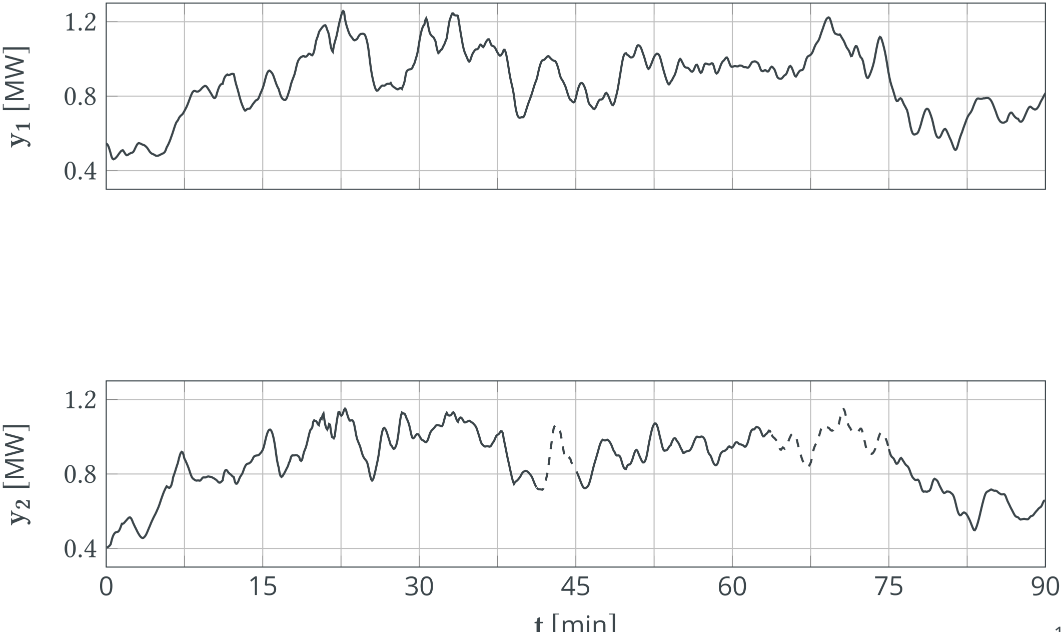
Hypothesis: A Bayesian graphical model



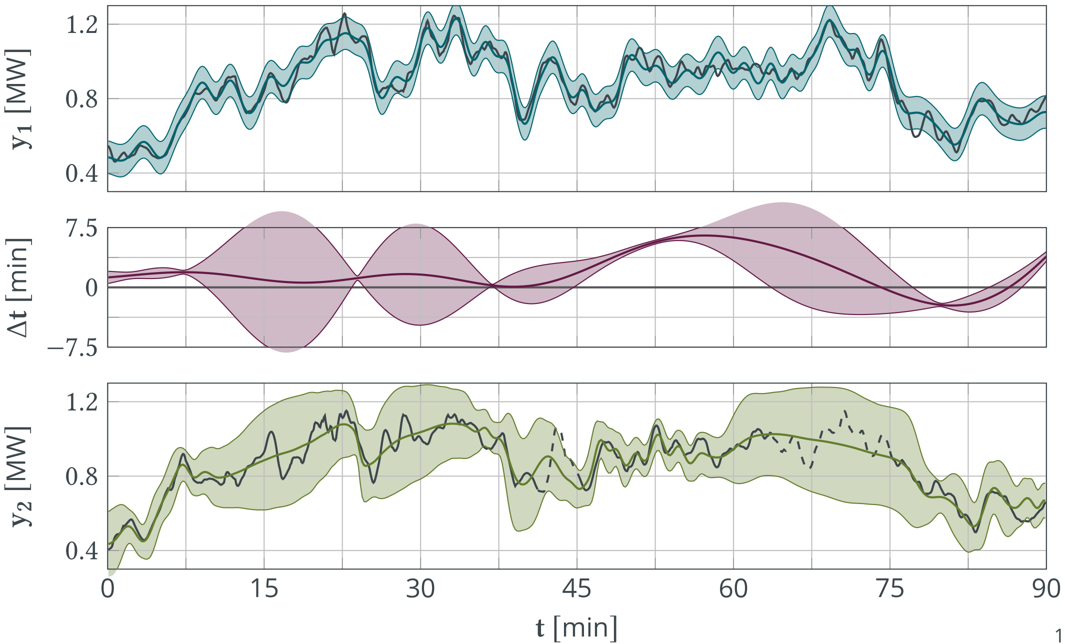
Hypothesis: A Bayesian graphical model



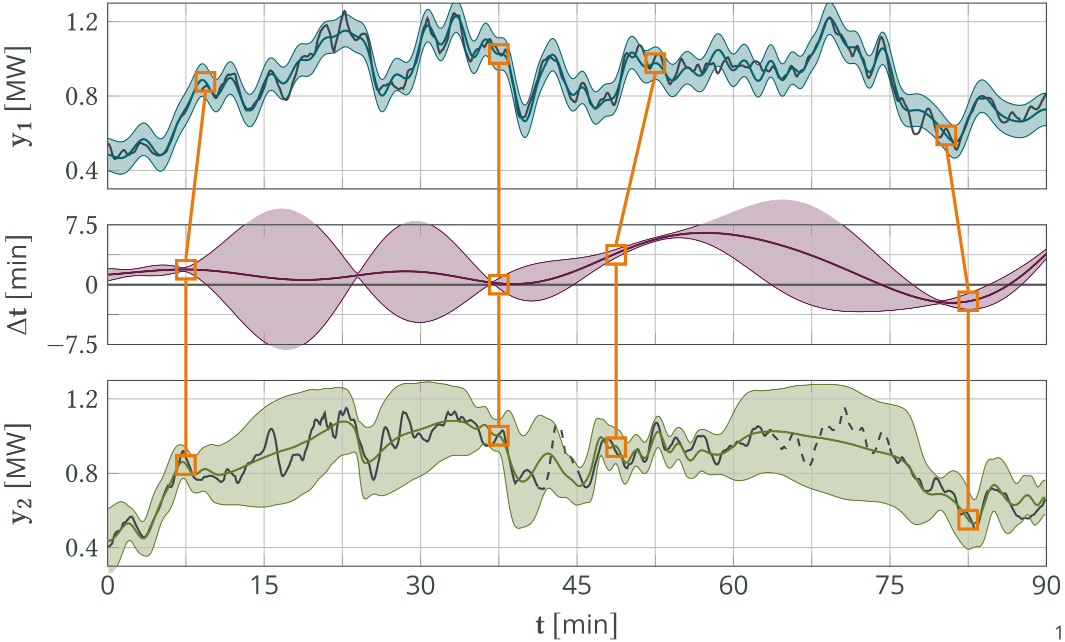
Posterior: Uncertain time alignment



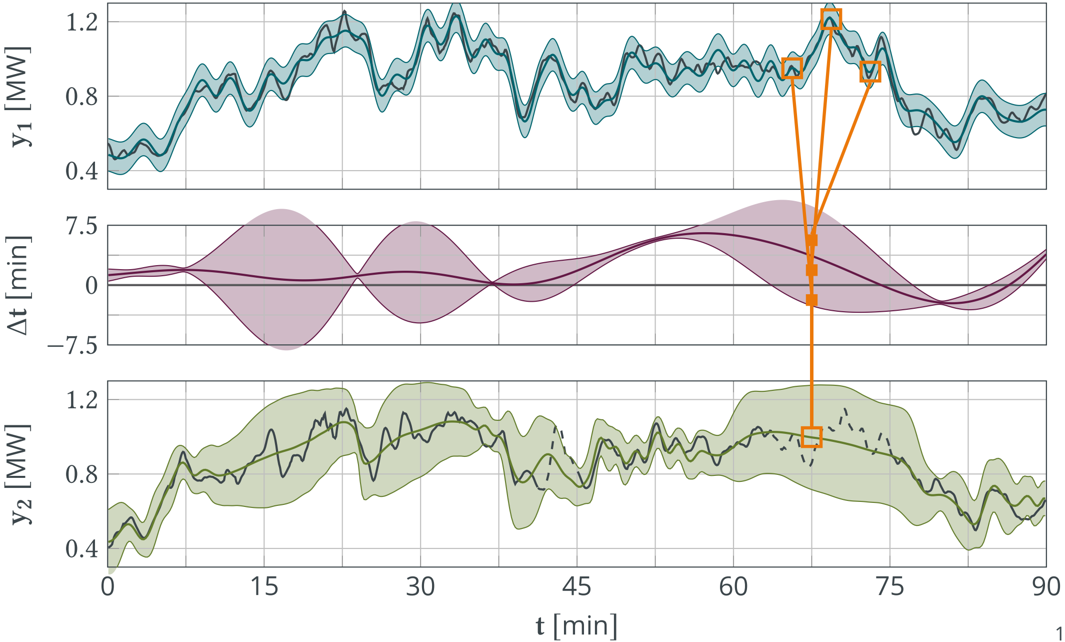
Posterior: Uncertain time alignment



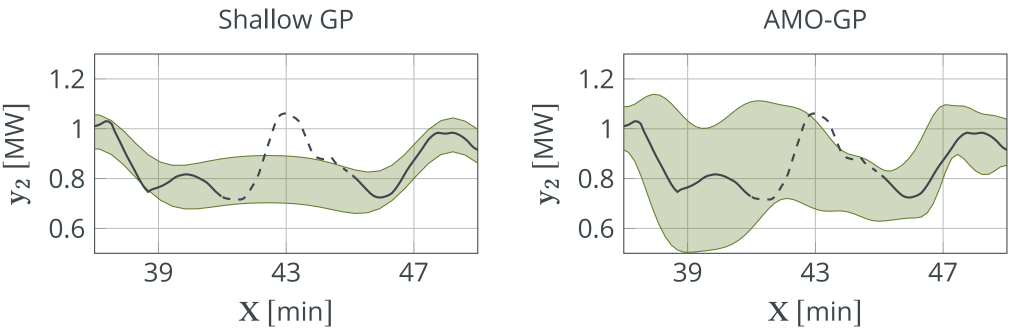
Posterior: Uncertain time alignment



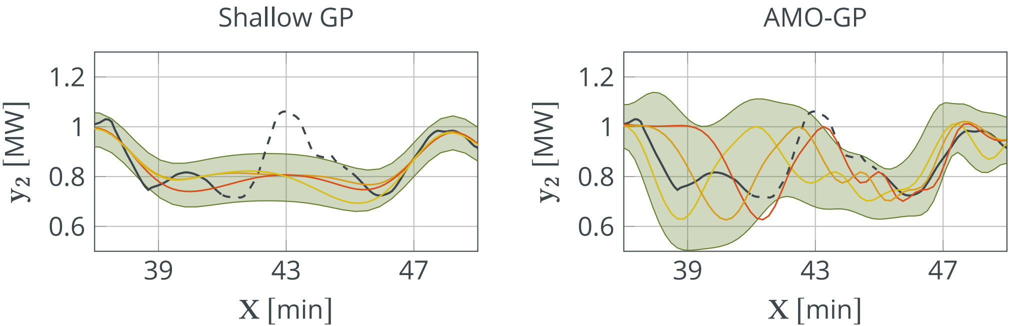
Posterior: Uncertain time alignment



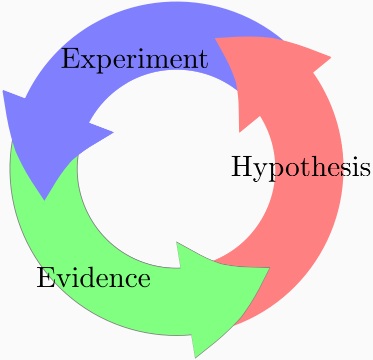
Comparing samples from the model



Comparing samples from the model

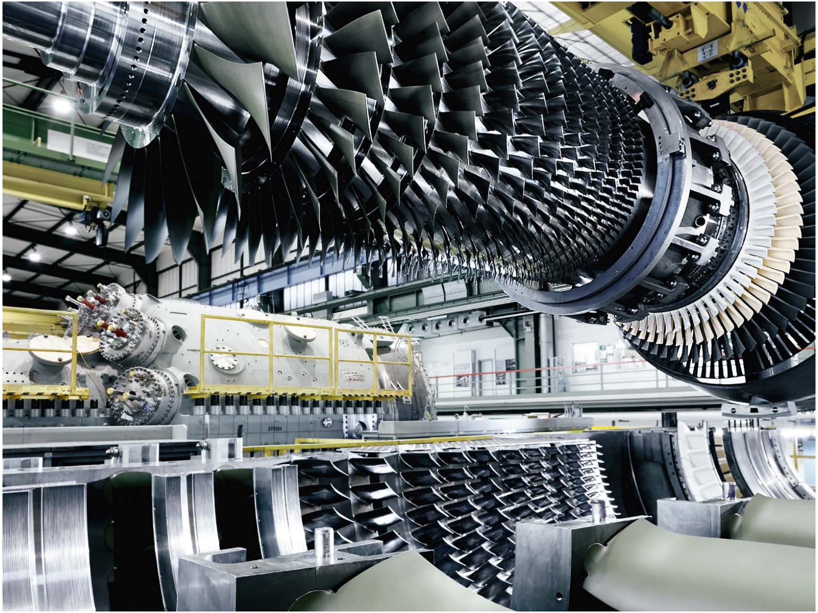


The Scientific Principle

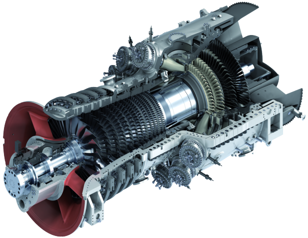


$$\text{Data} + \text{Model} \overbrace{\rightarrow}^{Compute} \text{Prediction}$$

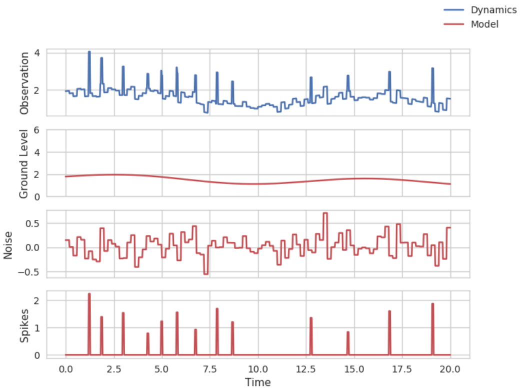
Gas turbines for power production



Data-Association model for gas turbines



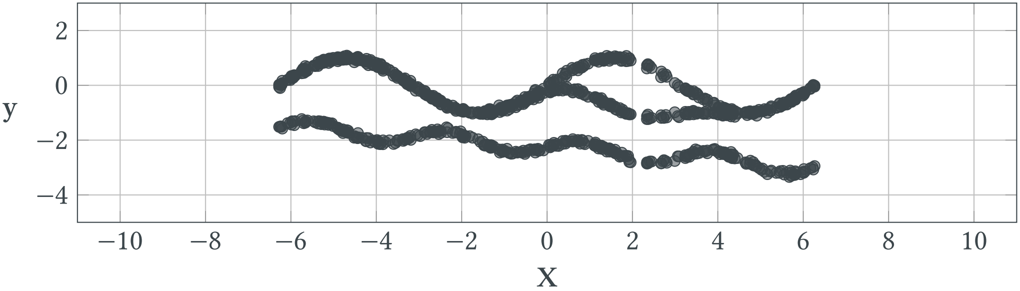
Siemens gas turbine



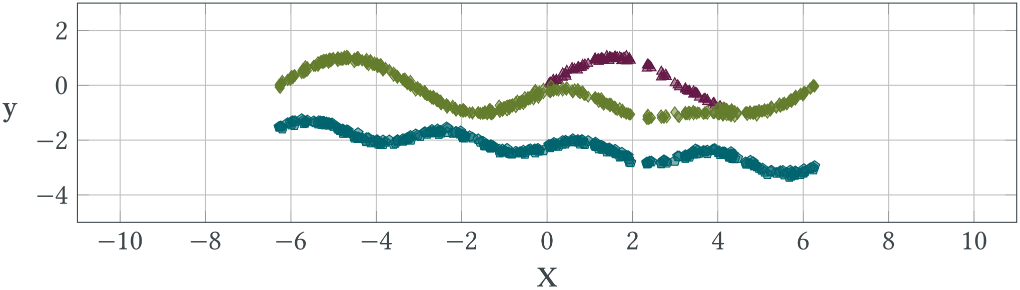
Combustion Dynamics

- Data from different operational regimes
- Robust inference for faulty sensors

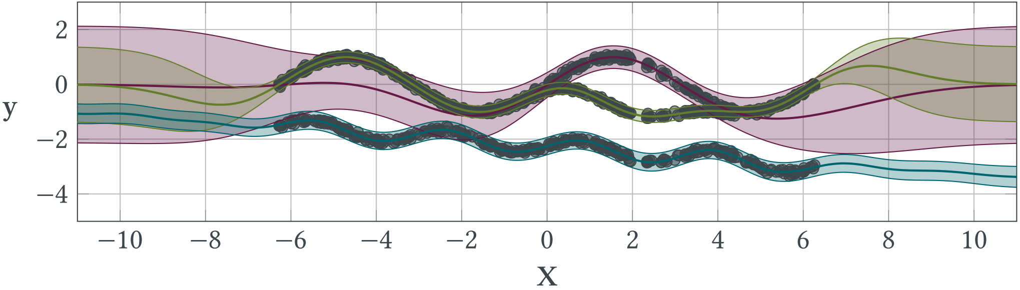
Multimodal data



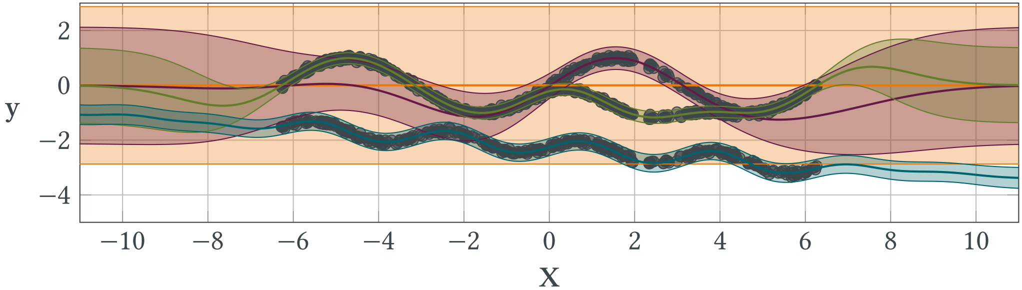
Multimodal data

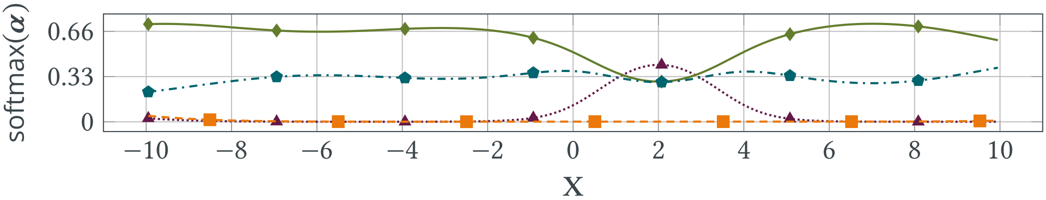


Multimodal data

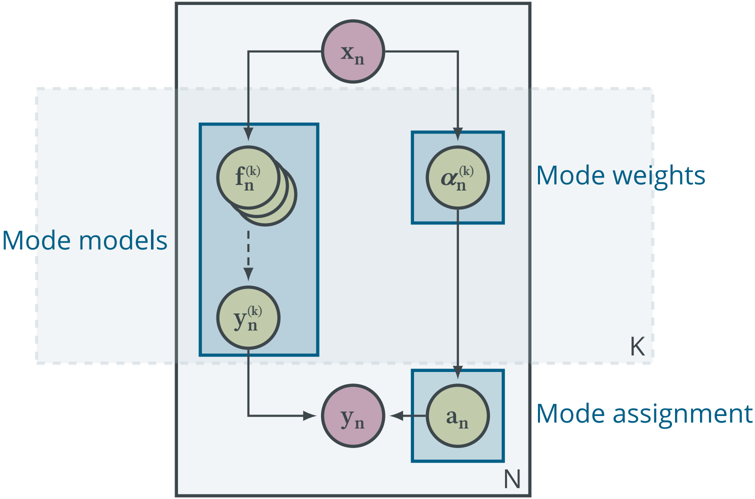


Multimodal data





Graphical Model of DAGP



The basic ML algorithm

Empirical risk minimization

- Approximate the **global true risk** wrt. loss ℓ

$$\mathrm{R}(f) := \int \ell(f(\mathbf{x}), \mathbf{y})\, \mathrm{p}(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}$$

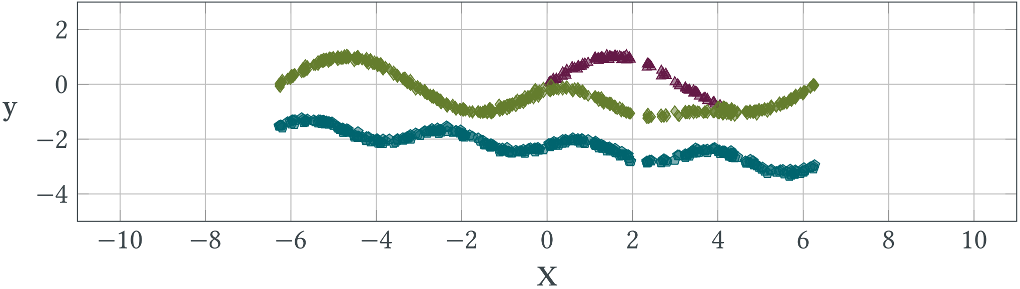
  with the **local empirical risk** in the available data

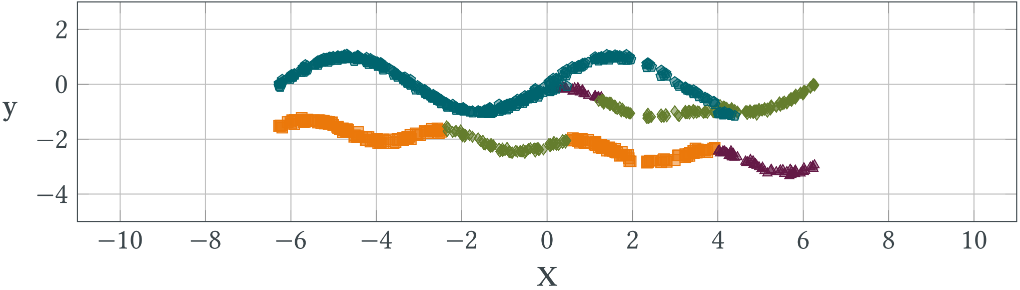
$$\mathrm{R}_{\mathrm{emp}}(f) := \frac{1}{N} \sum_{i=1}^{N} \ell(f(\mathbf{x}_i), \mathbf{y}_i)$$

- **Learning algorithm:** Choose a hypothesis space $\mathscr{H} \subseteq \mathscr{F}$ and use

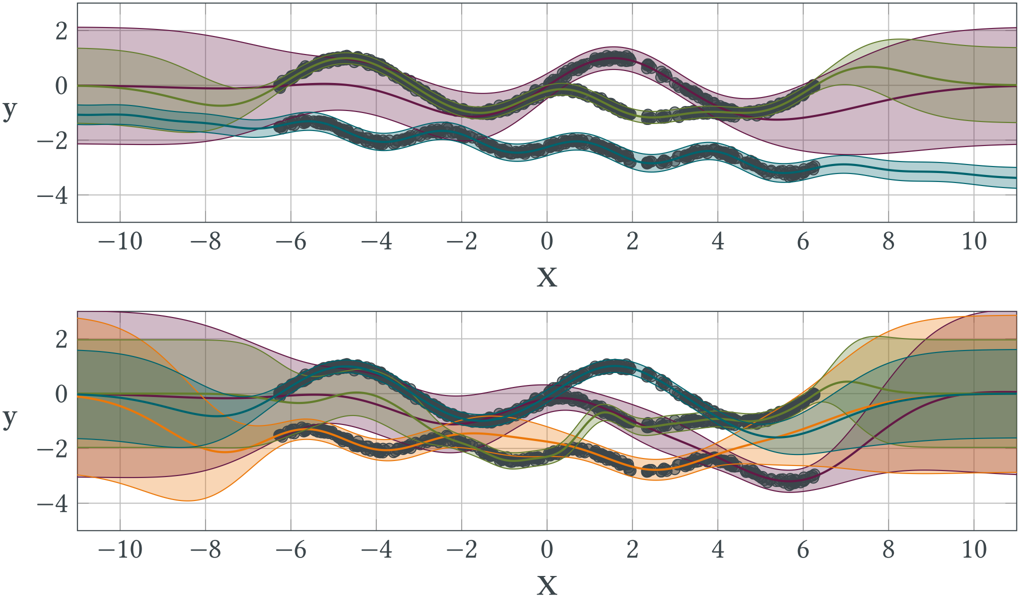
$$\hat{f} \in \underset{f \in \mathscr{H}}{\operatorname{argmin}}\, \mathrm{R}_{\mathrm{emp}}(f)$$

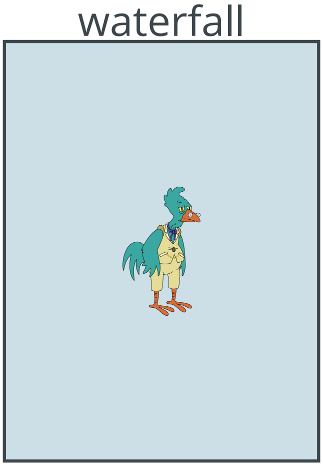
Multimodal data





Multimodal data





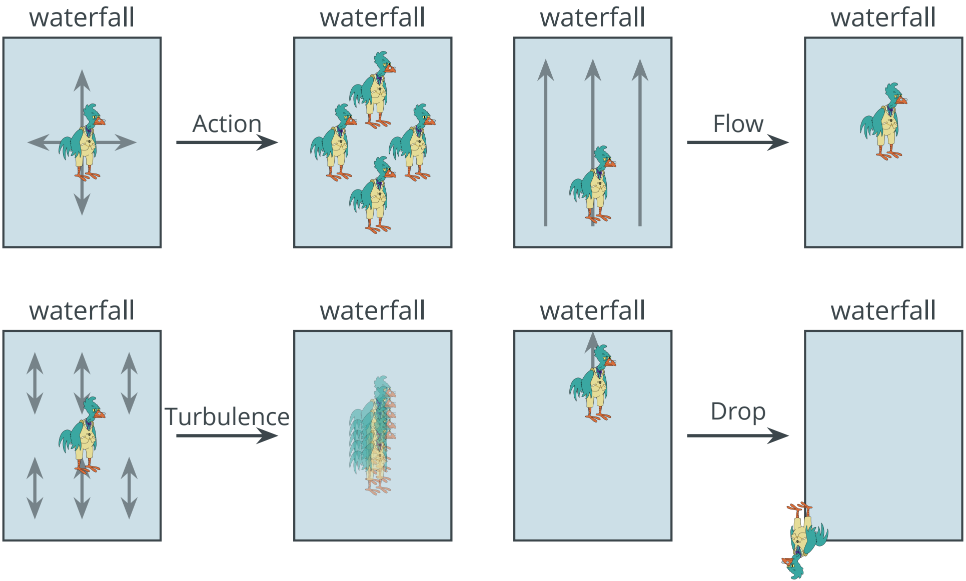
Dynamics Agent in a flowing river

Goal Get close to the waterfall

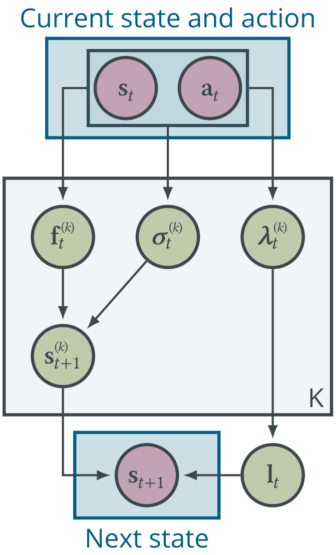
State (x, y)-position in $\mathbb{R}^2$

Action (x, y)-movement in $\mathbb{R}^2$

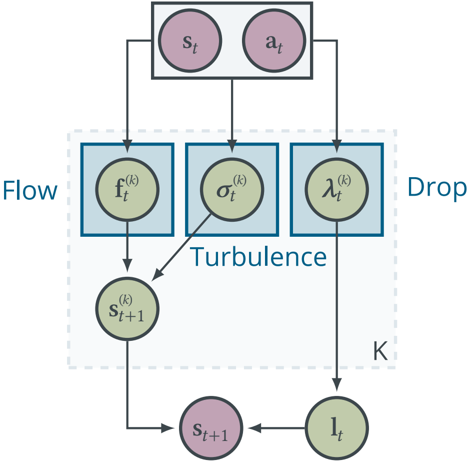
Wet-Chicken Benchmark



Graphical Model of DAGP



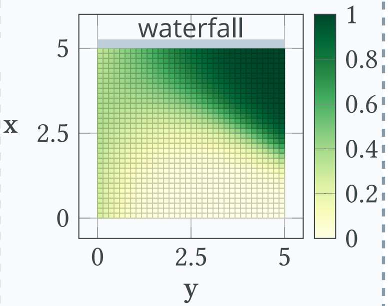
Graphical Model of DAGP



Multimodal System Dynamics

$$\mathrm{p}(\Delta\mathbf{x}_{t+1}) = \mathrm{p}(\Delta\mathbf{x}_{t+1}\,|\,\mathrm{drop}) \cdot \mathrm{p}(\mathrm{drop})$$
$$+\ \mathrm{p}(\Delta\mathbf{x}_{t+1}\,|\,\mathrm{no\ drop}) \cdot \mathrm{p}(\mathrm{no\ drop})$$

p(drop)

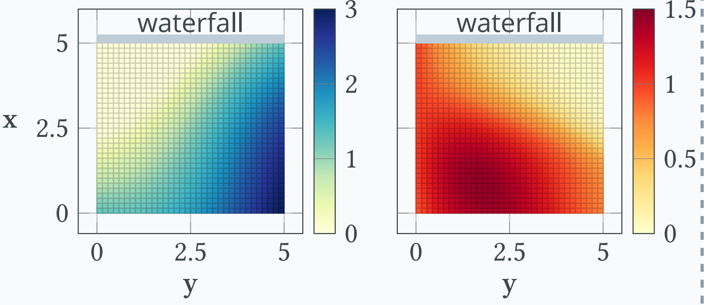


Flow

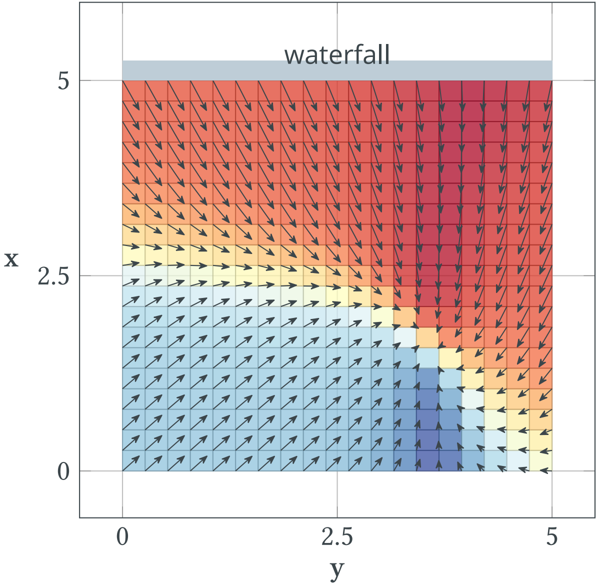
$\mathbb{E}[\Delta\mathbf{x}_{t+1}\,|\,\mathrm{no\ drop}]$

Turbulence

$\sigma[\Delta\mathbf{x}_{t+1}\,|\,\mathrm{no\ drop}]$



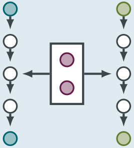
Wet-Chicken Policy



Summary

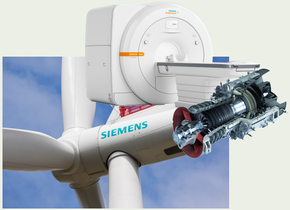
Scientific and industrial AI

- Models must stand up to scrutiny
- Knowledge is often hierarchical
- Enforce scientific plausibility



Subjectivity of models

- ML is great at explaining data
- But not all explanations are valid
- Beyond metrics, experts need to judge



Markus Kaiser — mrksr.de