ACCRUE: Accurate and Reliable Uncertainty Estimate in Deterministic Models

Enrico Camporeale (enrico.camporeale@noaa.gov)

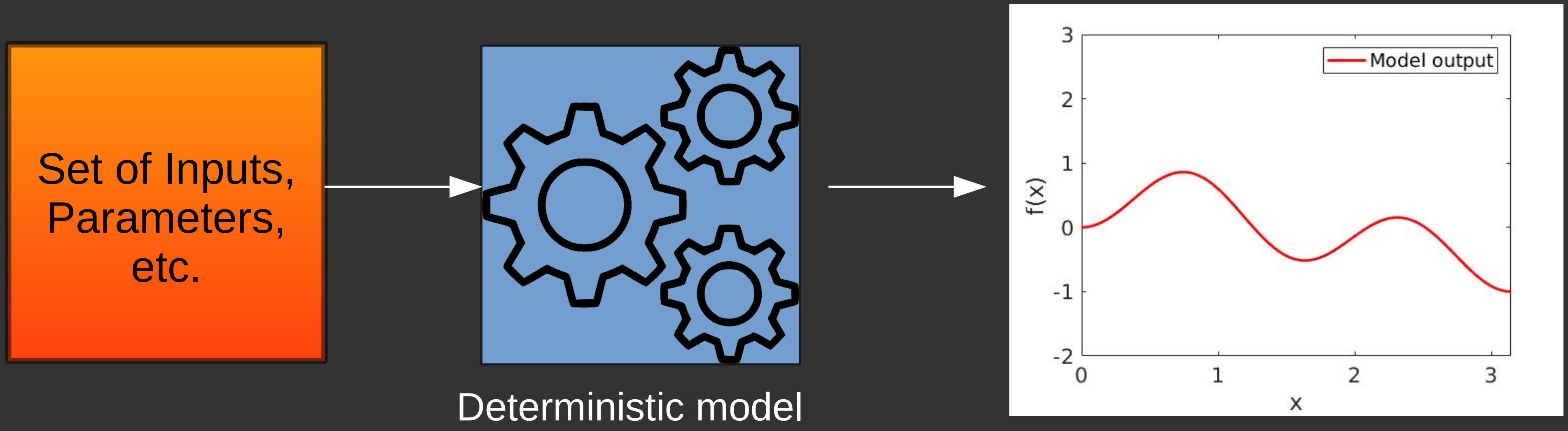
CIRES / CU Boulder & NOAA Space Weather Prediction Center

This project is supported by NASA under grant 80NSSC20K1580

University of Colorado Boulder

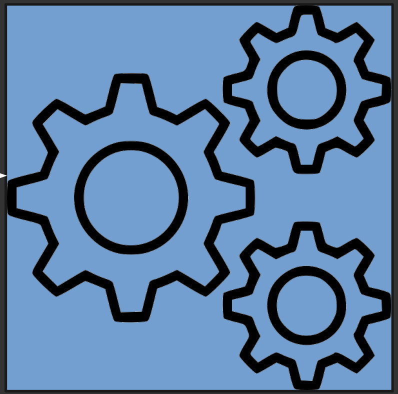
CIRES

Problem statement

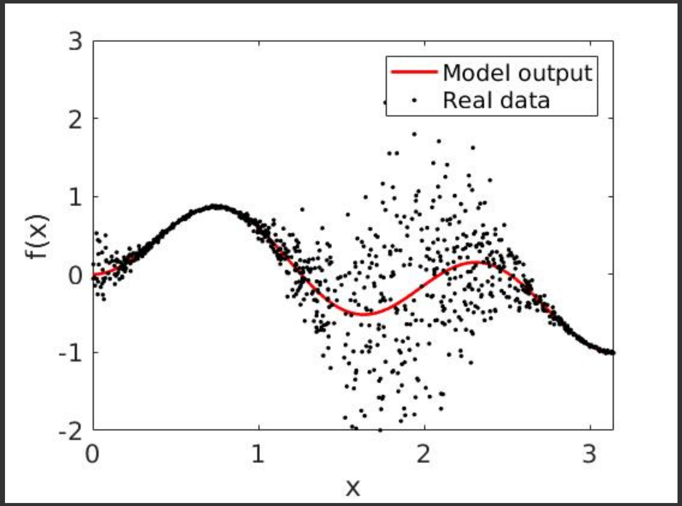


Problem statement

Set of Inputs, Parameters, etc.



Deterministic model + Real Data

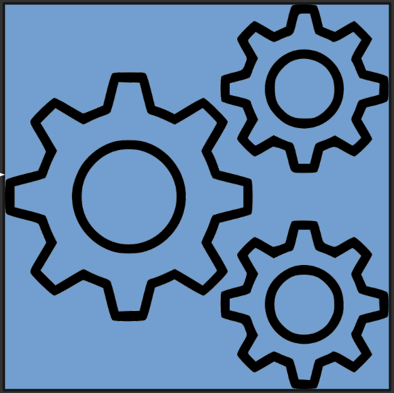


Uncertainties:

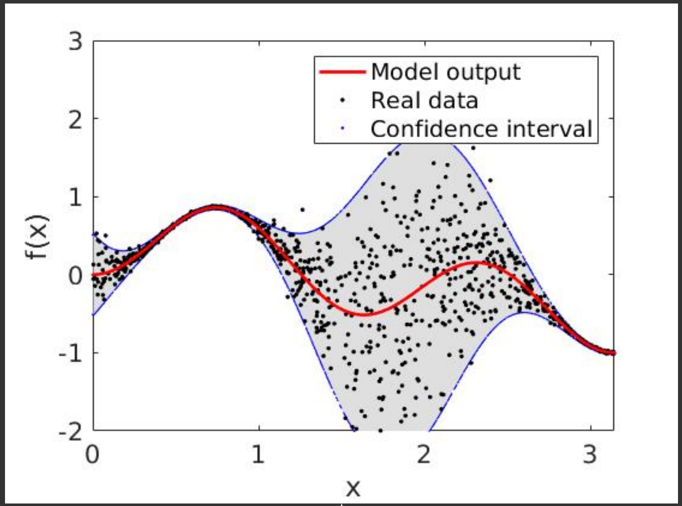
- Epistemic (do not know the parameters exactly)
- Aleatoric (physics that is not in the model)
- Algorithmic (numerical errors)

Problem statement

Set of Inputs, Parameters, etc.



Deterministic model + Real Data



Uncertainties:

- Epistemic (do not know the parameters exactly)
- Aleatoric (physics that is not in the model)
- Algorithmic (numerical errors)

?

Generate heteroskedastic uncertainty

Uncertainty Quantification in Machine Learning

- A relatively new research topic
- Parametric models:
  - Bayesian Neural Networks
  - MC Dropout
  - Ensemble techniques

The model output is typically not well-calibrated / reliable

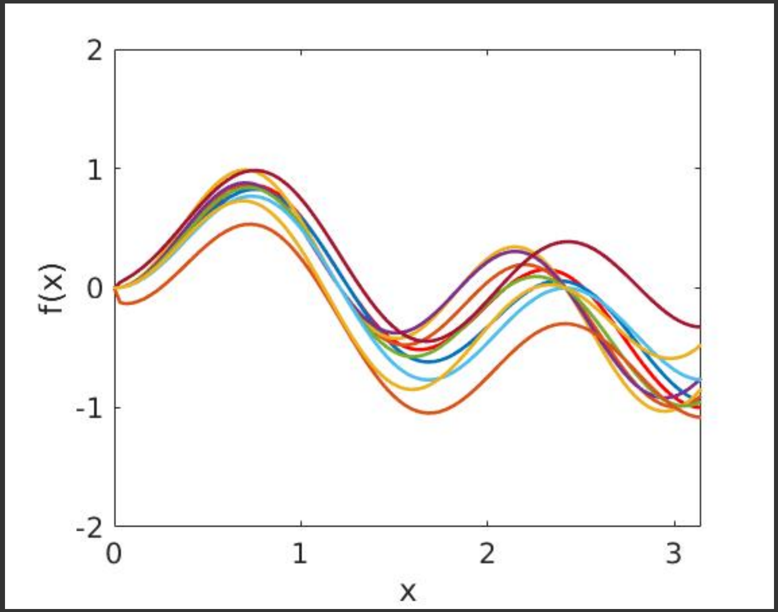
- Non-parametric models:
  - Gaussian Process

Algorithm does not scale well with large datasets

UQ in Physics-Based models

- The golden standard approach to estimate uncertainties based on a deterministic model is by running a

  Monte Carlo ensemble
  (e.g. by small perturbations of initial conditions)

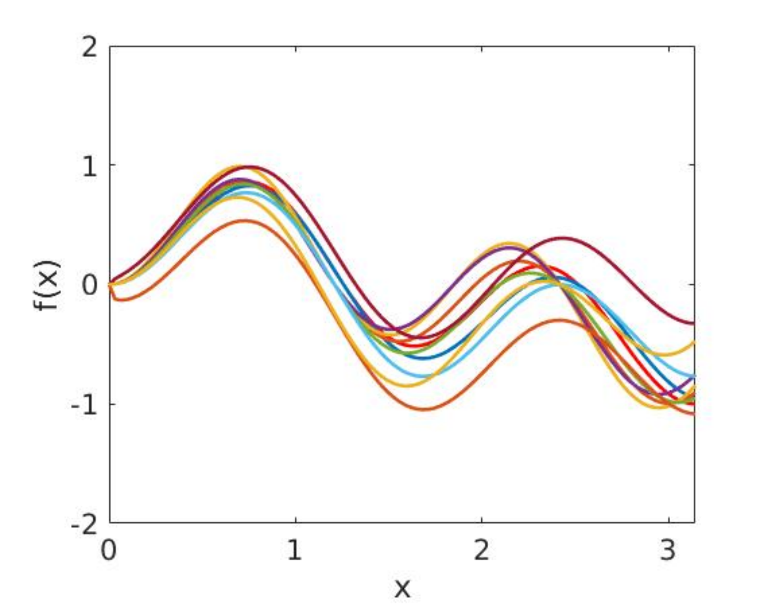


- This has two problems:
  - It requires many runs (expensive)
  -

UQ in Physics-Based models

- The golden standard approach to estimate uncertainties based on a deterministic model is by running a

  Monte Carlo ensemble
  (e.g. by small perturbations of initial conditions)



- This has two problems:
  - It requires many runs (expensive)
  - It requires to know what is the probability distribution of inputs and how to draw samples from it

Take home message

We have devised a method that:

- Estimates the uncertainties associated with single-point outputs generated by a deterministic model, in terms of Gaussian distributions;
- Ensures the optimal trade-off between accuracy and reliability;
- Does not need to run ensembles. It costs as much as training a neural network
- Code available:

  zenodo.1485608

International Journal for Uncertainty Quantification, 11(4):81–94 (2021)

ACCRUE: ACCURATE AND RELIABLE UNCERTAINTY ESTIMATE IN DETERMINISTIC MODELS

Enrico Camporeale[1,*] *& Algo Carè*[2]

[1]*University of Colorado, Boulder, Colorado, USA*

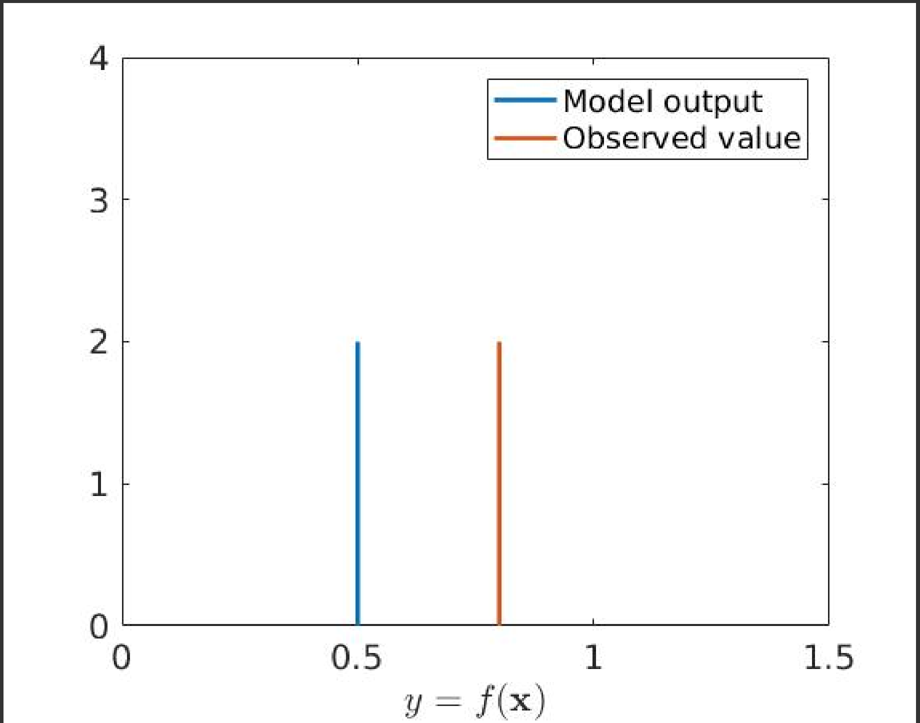
[2]*University of Brescia, Brescia, Italy*

Utilitarian Approach

Let us assume that for a single (multidimensional) input **x**, our model predicts an output y = f(x).

Blue line → Model output
Red line → Real (observed value)

<u>Working hypothesis:</u>
We want to use the model output ("oracle") as the mean of a Gaussian distribution that is interpreted as a probabilistic forecast.



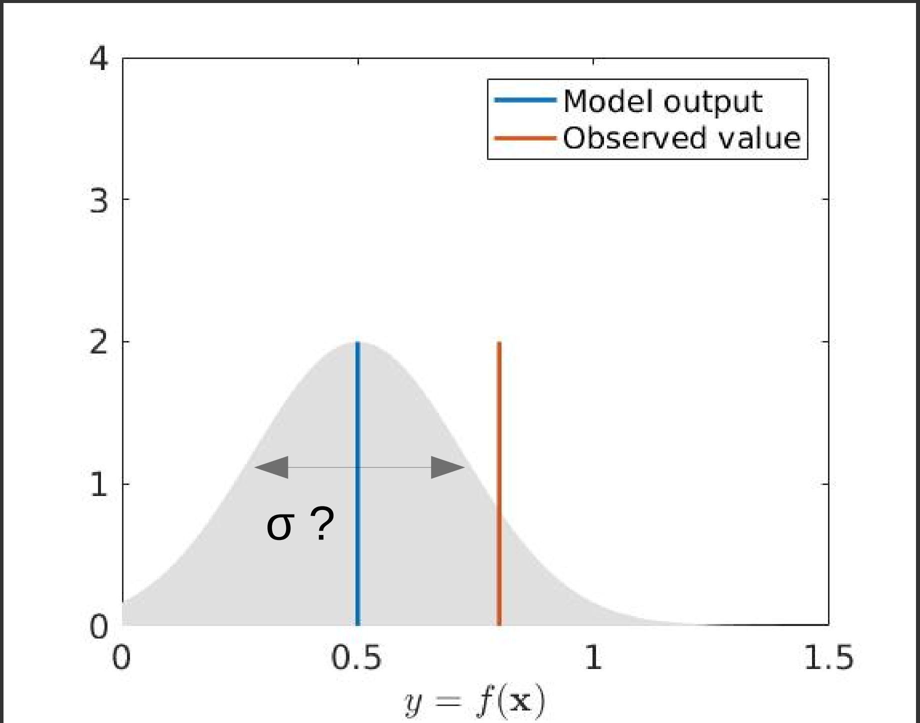
Utilitarian Approach

Let us assume that for a single (multidimensional) input **x**, our model predicts an output y = f(x).

Blue line → Model output
Red line → Real (observed value)

Working hypothesis:
We want to use the model output ("oracle") as the mean of a Gaussian distribution that is interpreted as a probabilistic forecast.

What is the optimal width of a Gaussian forecast?
The one that optimizes BOTH accuracy and reliability

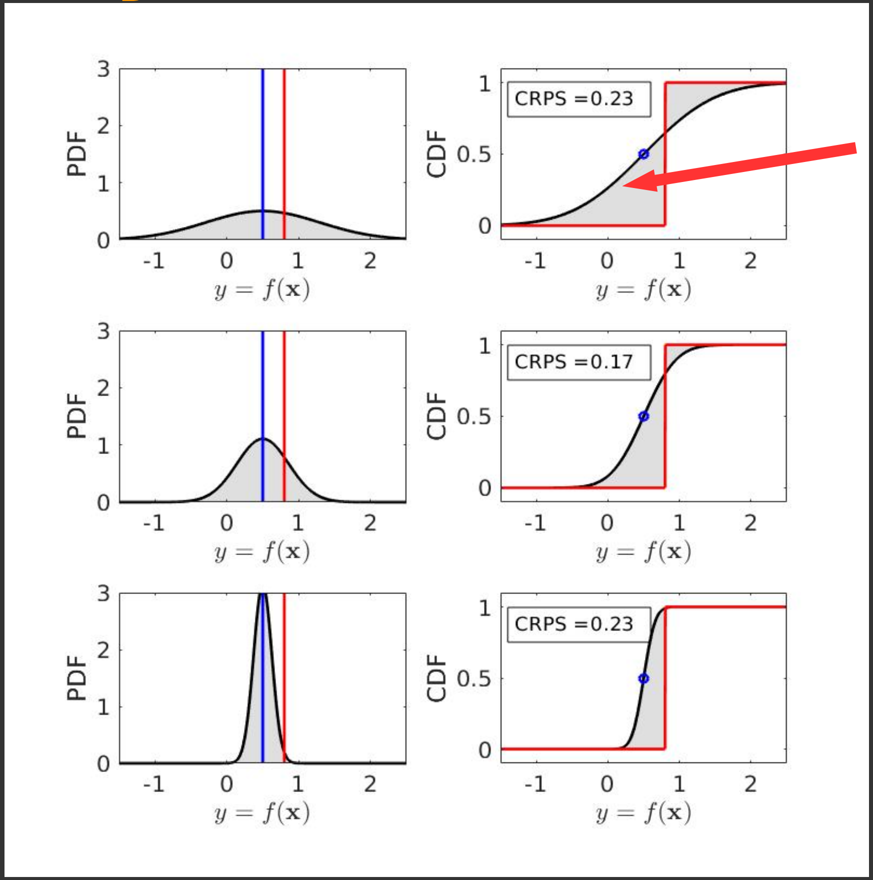


Accuracy: Continuous Rank Probability Score

- **CRPS is a generalization of Brier score**
- **It has a simple graphical interpretation**
- **CRPS = 0 for perfect forecast**
- **CRPS =** $\int \left(C(y) - H(\hat{y})\right)^2 dy$

CDF

Step function

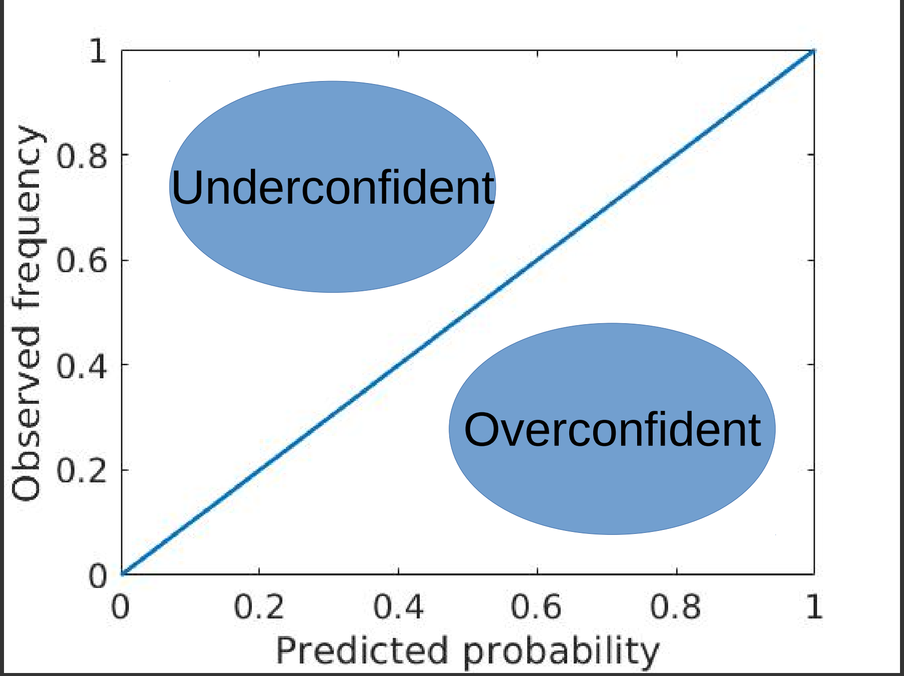


CRPS proportional to this area

Reliability diagram

Reliability is the property of a probabilistic model that measures its **statistical consistency with observations.**

For example, for forecasts of discrete events, the reliability measures if an event occurs on average with frequency *p*, when it has been predicted to occur with probability *p*.



Reliability diagram

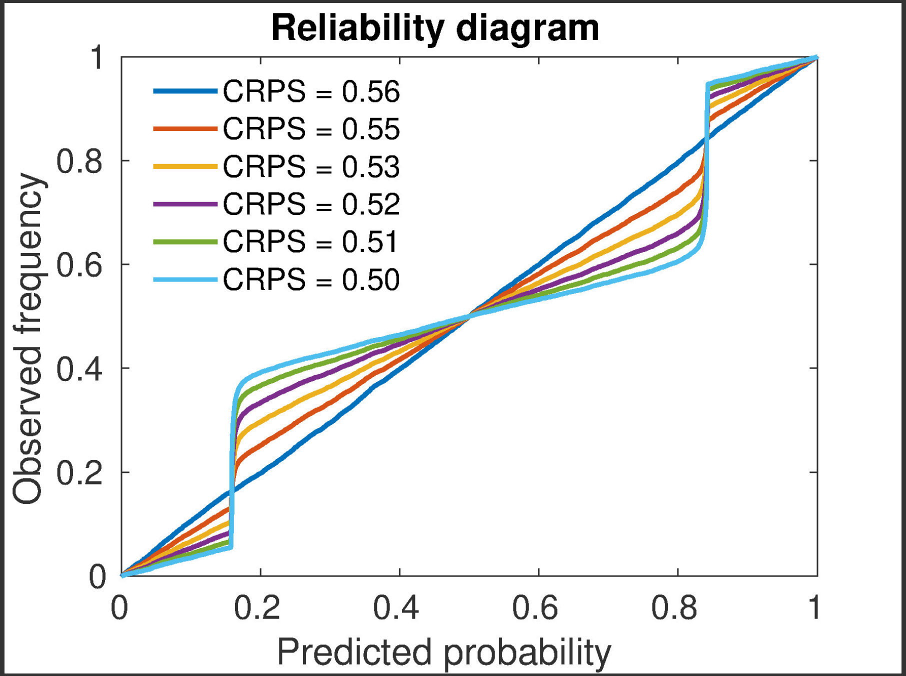
The values of σ that minimize CRPS can be derived analytically:

$$\sigma^2_{opt} \sim \varepsilon^2$$

The more we minimize CRPS, the worse reliability we get.

Mathematical proof is straightforward.

See arXiv:1811.12692



Two-objective cost function

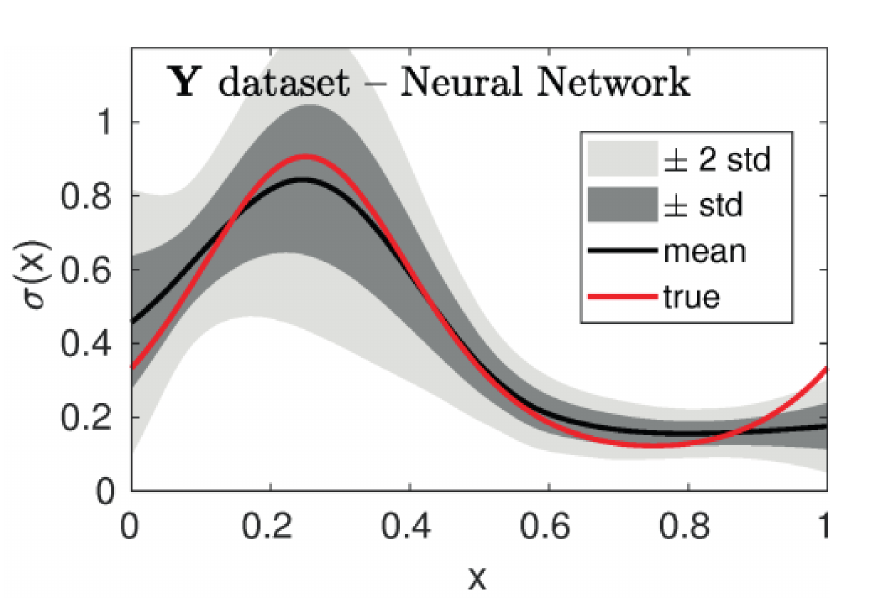
- This is a two-objective optimization problem, because <u>reliability</u> and <u>accuracy</u> are competing objectives.
- We define the Accuracy-Reliability (AR) cost function:

$$AR = \underbrace{CRPS}_{\text{Accuracy}} + \beta * \underbrace{RS}_{\text{Reliability}}$$

- Accuracy and Reliability cannot both be minimized simultaneously
- We have to find the best trade-off

The ACCRUE Method

- Take a sample of N errors ε (difference between model output and observed values) and the corresponding model inputs **x**
- We define as optimal standard deviation σ the one that optimizes the Accuracy-Reliability cost function (that has an analytical expression)
- We also want to have a way of generating a smooth function σ(**x**) for any value of **x**
- We use a neural network that takes **x** as input and produces σ(**x**) as output by minimizing AR cost function.



Machine Learning benchmark

TABLE 1: Comparison between different methods on several multidimensional datasets. Median values are reported, calculated over 50 runs. Confidence intervals represent one standard deviation. Best values are in bold

Method / Score			CRPS	RECAL CRPS	KM	ACCRUE
Dataset	**Size**	**Dim.**				
Boston Housing	506	13	0.25 ± 0.05	0.25 ± 0.04	0.25 ± 0.03	$\mathbf{0.23 \pm 0.04}$
Concrete	1030	8	0.22 ± 0.03	0.23 ± 0.13	0.26 ± 0.02	$\mathbf{0.21 \pm 0.03}$
Energy	768	8	0.059 ± 0.03	0.056 ± 0.03	0.087 ± 0.01	$\mathbf{0.052 \pm 0.01}$
Kin8nm	8192	8	0.17 ± 0.005	0.16 ± 0.01	0.24 ± 0.005	$\mathbf{0.16 \pm 0.005}$
Power plant	9568	4	0.13 ± 0.003	0.13 ± 0.05	0.15 ± 0.002	$\mathbf{0.12 \pm 0.01}$
Protein	45,730	9	0.38 ± 0.02	0.47 ± 0.13	0.40 ± 0.007	$\mathbf{0.37 \pm 0.02}$
Wine	1599	11	0.48 ± 0.03	0.50 ± 0.29	$\mathbf{0.46 \pm 0.02}$	0.48 ± 0.06
Yacht	308	6	0.06 ± 0.08	0.06 ± 0.02	0.19 ± 0.02	$\mathbf{0.06 \pm 0.02}$
Score			**Cal. err. (%)**			
Dataset	**Size**	**Dim.**				
Boston Housing	506	13	26.2 ± 7.9	20.6 ± 5.5	17.5 ± 3.7	$\mathbf{16.7 \pm 5.9}$
Concrete	1030	8	22.6 ± 5.8	14.4 ± 3.8	22.1 ± 3.0	$\mathbf{11.5 \pm 3.9}$
Energy	768	8	29.3 ± 8.9	29.2 ± 8.0	28.3 ± 2.8	$\mathbf{13.0 \pm 6.5}$
Kin8nm	8192	8	15.9 ± 1.28	8.3 ± 1.30	25.5 ± 0.5	$\mathbf{5.8 \pm 1.28}$
Power plant	9568	4	12.5 ± 1.4	3.4 ± 0.9	16.1 ± 0.8	$\mathbf{2.6 \pm 0.8}$
Protein	45,730	9	13.1 ± 0.8	$\mathbf{5.0 \pm 0.9}$	10.6 ± 0.9	5.4 ± 0.88
Wine	1599	11	16.0 ± 3.7	$\mathbf{7.9 \pm 2.0}$	8.0 ± 2.4	8.3 ± 2.4
Yacht	308	6	26.0 ± 9.4	24.3 ± 13.5	36.6 ± 3.0	$\mathbf{19.5 \pm 8.5}$

Take home message

We have devised a method that:

- Estimates the uncertainties associated with single-point outputs generated by a deterministic model, in terms of Gaussian distributions;
- Ensures the optimal trade-off between accuracy and reliability;
- Does not need to run ensembles. It costs as much as training a neural network
- Code available:

  zenodo.1485608


International Journal for Uncertainty Quantification, 11(4):81–94 (2021)

ACCRUE: ACCURATE AND RELIABLE UNCERTAINTY ESTIMATE IN DETERMINISTIC MODELS

Enrico Camporeale[1,] & Algo Carè[2]*

[1]*University of Colorado, Boulder, Colorado, USA*
[2]*University of Brescia, Brescia, Italy*